DOCUMENT RESUME

ED 157 911 TM 007 220

AUTHOR	Gephart, William J.; And Others
TITLE	Evaluation in the Affective Domain. NSPER: 76.
INSTITUTION	Phi Delta Kappa, Bloomington, Ind.
PUB DATE	76
NOTE	191p.; Papers presented at the National Symposium for Professors of Educational Research (Memphis, Tennessee; Baltimore, Maryland; and Phoenix, Arizona, 1976)
AVAILABLE FROM	Phi Delta Kappa, Center on Evaluation, Development, and Research, Eighth Street and Union Avenue, Box 789, Bloomington, Indiana 47401 ($3.50)
EDRS PRICE	MF-$0.83 HC-$10.03 Plus Postage.
DESCRIPTORS	*Affective Objectives; *Affective Tests; Attitude Tests; Classroom Environment; Evaluation Methods; *Humanistic Education; Learning Motivation; *Measurement Techniques; Physiology; School Environment; *Self Concept; *Student Attitudes; Theories

ABSTRACT

The National Symposium for Professors of Educational Research for 1976 focused on two topics: the nature of affect, and principles and guidelines for measuring individual affect and learning environment. This document contains five major papers presented at the conference. The first paper contrasted the physiological and emotional concept of affect as a feeling-state of unusual intensity, with the educational concept of affect as the link between cognition and behavior, emphasizing the perceptual aspects of affect and self concept, self esteem, and the learning climate in the schools. The second paper discussed affect from a biochemical perspective. The third paper emphasized the relation between individual affect and dimensions of educational environments, and described a number of instruments for evaluating school or class climate and studies relating them to achievement test scores. The fourth paper described a self-concept theory of motivation and learning with emphasis on feelings of personal worth, ability to cope, ability to express, and ability to make choices. The last paper discussed evaluation as measurement, decision making, and judgment; described the affective domain as psysiological and psycho-social including values, emotions, and perceptions; and commended the acceptance of an affective epistemology. (CTM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Juliette Wilson
CEDR

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC) AND
USERS OF THE ERIC SYSTEM."

ED157911

TM007 320

Evaluation in the Affective Domain

William J. Gephart
Robert B. Ingle
and
Frederick J. Marshall
Editors

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

A CEDR Monograph
PHI DELTA KAPPA

NATIONAL SYMPOSIUM FOR PROFESSORS OF EDUCATIONAL RESEARCH

NSPER: 67 *Alternative Approaches for Teaching the Four Basic Research Areas* -- University of Maryland

NSPER: 68 *Instructional Materials Development for Research Instruction* – University of Colorado

NSPER: 69 *Research Techniques from Selected Disciplines that are Possibly Applicable to Education* – University of Chicago

NSPER: 70 *Introductory Research Course: Content, Technique and Resources* – Washington University, St. Louis

NSPER: 71 *Graduate Measurement Instruction* – University of Wisconsin/Milwaukee, University of Florida, University of Arizona

NSPER: 72 *Similarities and Differences in the Research and Evaluation Processes* – Michigan State University, San Francisco, University of North Carolina

NSPER: 73 *Teaching About Evaluation* – University of Wisconsin/Milwaukee, SUNY/Albany, UCLA, Chicago

NSPER: 74 *The Evaluation of Teaching* – Denver, Seattle, Atlanta

NSPER: 75 *The Evaluation of Administrative Performance: Parameters, Problems and Practices* – Ohio State University, University of Oregon, Texas Christian University

NSPER: 76 *Evaluation in the Affective Domain* – Memphis, Baltimore, Phoenix

PREFACE

The National Symposium for Professors of Educational Research (NSPER) was conceived in the mid-60's as a conference for dealing with concerns of individuals with responsibility for teaching research training programs. Specifically, two objectives were behind the creation of NSPER: 1) to provide an opportunity for people who teach about research and related skills to exchange information about problems, materials and techniques in their instructional assignments, and 2) to help these individuals obtain new information from leaders in the field.

Following the enthusiastic response of the participants to the first conference in 1967 which focused on alternative approaches for teaching the four basic research areas (introduction, measurement, statistics and advanced design), the symposium developed into an annual series. Topics have ranged from the design of research and measurement courses to evaluation techniques. These topics are chosen in light of current concerns in the field of educational research.

NSPER: 76 represents the end of the first decade of this conference series. Fittingly, perhaps, the topic of this symposium is the definition and measurement of affect as a component of the learning environment. The presentations of the major speakers at the three sessions – Memphis, Baltimore and Phoenix – have been collected and are provided in this report. It is hoped that this report adequately reflects the significant ideas and issues that were raised and discussed at the three conferences.

WJG
RBI
FJM

TABLE OF CONTENTS

PREFACE iii

INTRODUCTION 3

NSPER: 76 – MEMPHIS

Affect And Its Measurement
Robert E. Bills. 7

NSPER: 76 – BALTIMORE

Affect: A Biochemical Perspective
David E. Price, M.D.. 55

The Assessment of Affect: Nomothetic and Ideographic
David A. Payne 61

NSPER: 76 – PHOENIX

Affective Measurement And The Self Concept
Walcott H. Beatty 119

NSPER: 76 – CO-DIRECTORS

Evaluation And The Affective Domain
William J. Gephart & Robert B. Ingle 165

INTRODUCTION

There has been increasing clamor for educational accountability in recent years. Parents and professionals alike are asking what the schools are doing to prepare children to meet future demands. More importantly, the questions are not only what the educational system plans on doing but how effective it is in accomplishing these goals.

In the certification of accountability of educational programs, broad new questions are also being posed. No longer is data on achievement sufficient. Educators are now being asked questions about attitudes, feelings, interests, values, etc. Such questions grow in importance as schools take on goals such as: "Develop pride in work and feelings of self worth," "Develop skills in interpersonal relationships," and "Establish a personal set of beliefs for daily living."

Answering questions about attitudes, interests, values or feelings is not easy. In the past, educational traditions, technology and practices have reflected a tendency to relegate these questions to a lesser status in comparison to questions concerning cognitive growth and development. This report from NSPER: 76, however, focuses on measurement undertaken to answer such questions. Two topics provide some structure for that focus: a) the nature of affect, and b) principles and guidelines for measuring individual affect and learning environment.

Three specific objectives were outlined for NSPER: 76. They were:

1. Participants will be exposed to, discuss and critique presentations which delineate:
   a) The nature of affect.
   b) The measurement of individual affect.
   c) The measurement of affect as a component of the learning environment.
2. Participants, working in small groups, will identify and rank in priority-order the principles and guidelines which should be considered in measuring affect on an individual and/or group basis.
3. Participants will use each other as resources through the exchange of information on problems encountered in measuring affect in specific settings.

Although the structure of the three sessions were identical, the content and affective climate of the conference varied across the sites. The major speakers addressed the issues of the nature and measurement of affect, yet provided their own unique perspective and drew upon their own individual experiences. The emotional tone or affective environment in each of the sessions seemed also to be heavily influenced by the opening dramatizations performed by school children from each of the three communities. The enthusiastic portrayals of the forms of affect by the children from Campus Elementary School in Memphis, Edmondson Senior High School in Baltimore and Central High in Phoenix provided a stimulating backdrop for the latter activities of the participants at the three sessions.

Unfortunately, it is near impossible to capture in words the richness of the experience which followed from the viewing of the various vignettes presented by the children. What this report documents, therefore, are simply the major addresses from the conferences. Contained in these papers is evidence of how difficult it is to adequately define or operationalize the concept of affect. It is most informative, however, to see how leaders in the field have come to grips with this definition and proceed to deal with issues involved in the measurement of affective components in the education process.

NSPER: 76

Memphis Presentation

Dr. Robert E. Bills

AFFECT AND ITS MEASUREMENT

Robert E. Bills
University of Alabama

When Bill Gephart called last spring and asked if I would work with this conference in the area of "Evaluation in the Affective Domain," I was delighted. But my delight turned to dismay when he stated in a letter at a later date, that in the first presentation, "We expect that you would start by stating what the concepts 'affect' and 'affective behaviors' mean to you." As you will see, that task is a difficult one since there is little agreement about the nature of the concept called affect, and the agreement which exists relates only tangentially to variables such as values, attitudes, self concepts, and others which are of concern to educators.

WHAT IS AFFECT AND WHY THE CONCERN?

What I want to do in this first presentation is to show that many educators and some psychologists use the term affect to describe an area of human existence whose nature is not clear to us. We know that education must be concerned with more than cognition and we have used the term affect to label some of the additional factors but the label is neither descriptive nor is it definitive. I would like also to talk about how people have attempted to solve the problem of defining affect and tomorrow to talk about problems involved in assessing what is purported to be affect and in helping schools improve in the so-called affective area.

Let me go back a short distance in the history of psychology to describe the affect problem, to show its lack of definition, and to establish that when educators and psychologists talk about affect they are not necessarily talking about the same thing. To clarify this last point, I suggest that psychologists who are interested in studying affect as a phenomenon see affect as a biological process, while educators and psychologists who are interested in helping people to change or to grow see affect as a cognitive type of process.

In 1928, Schneider wrote,

> In the psychology of feeling, although there is an every-growing accumulation of material, there is no satisfactory system and classification. Investigators are interested chiefly in the physiological foundation of the expressions of emotion, whereas the experiences themselves are neglected (p. 475).

In most respects, the statement is valid today and it points to the experiential and physiological horns of our definitional dilemma. The statement is a most interesting one in that it recognizes the content of the experience as a factor which should be of concern but which is neglected because of concentration on the physiology of emotion. Today, most psychologists who are interested in emotion view it as a physiological state, and since positive physiological states of emotion cannot be demonstrated, the psychology of emotion is the psychology of fear and anger. Little is said about positive emotions such as love, and the experiential aspects of emotion or affect are ignored.

Definitions of affect over the past 50 years differ little from each other. In 1934, Warren gave this definition, "Affect = . . . the dynamic and essential constituent of emotion . . . [It] is an inclusive term, used to denote any variety of emotional experience or emotional concomitants, especially a strong emotion, and also moods" (p. 7). The equation of affect and emotion is shown also in Duncker and Watt's (1938) *German-English Dictionary of Psychological and Psychoanalytical Terms* which translates the German word "affekt" to " . . . affect, emotion" and in Beigel's (1971) *Dictionary of Psychology and Related Fields* which also translates "affekt" to " . . . affect, emotion."

In 1958, English and English described affect as " . . . a class name for feeling, emotion, mood, temperament . . . Many contemporary theorists hold that there is no separate affect-state, but only an affective or feeling aspect of a cognitive state or process . . . "(p. 15). Thus, it can be seen that some psychologists recognized a non-physiological definition of affect but their number became smaller as psychology moved during the past 25 years toward a definition of itself as the study of behavior by which is meant the study of overt, observable, measurable, behavioral acts. This movement left little room for affect as an important component of human existence.

English and English did not overlook the more traditional definitions entirely. Thus, they stated that affect "practically = an emotion . . . . the dynamic or essential quality of an emotion; the energy of an emotion" (1958, p. 15). In so doing, they negated the most important part of their first definition, that affect may be a feeling aspect of a cognitive state or process.

Wolman (1973) illustrates the points I am attempting to make which are that: 1) conceptual confusion is present in most definitions of affect, 2) much of psychology has attempted to equate affect with emotion as a means of dealing with affect, 3) there has been little, if any, change in definitions in the past 50 years, and 4) affect can be either a perceptual or a physiological state, although it is more frequently defined as physiological rather than perceptual. In his 1973 dictionary, Wolman defines affect as: "1. A class name given to feelings, emotions, or dispositions as a mode of mental functioning. 2. The name given to specific emotions or feelings. 3. . . . A state generated when motivated action becomes unfeasible" (p. 11).

An excellent illustration of my point regarding confusion which appears in definitions of affect can be found in a 1972 description by Eysenck:

> **Affect**. A term that is not defined uniformly. In general, it is used to characterize a feeling state of particular intensity. Sometimes an 'affect' is characterized as a state brought about by actions almost wholly devoid of intentional control in accordance with moral and objective viewpoints. The term is also found in the . . . [psychological] . . . literature as practically synonymous with 'emotion' in certain senses (p. 28).

Formal definitions also support my point regarding the lack of relevance of the terms affect or emotion for education; none encompasses what I believe educators are talking about when they talk about affect. To clarify this statement let me call attention to the fact that with the exception of the term "moral viewpoint" used by Eysenck, none of the definitions I have presented deals with terms such as values, ideals, attitudes, self concepts, morals, character, and others which concern educators. All of these considerations lead me to suggest that we probably need a new term or viewpoint which does not equate affect and emotion and which is broad enough to encompass educators' concerns for the nature of the inner experience of people which seems so directly related to the quality of their behavior.

When I examine instruments which purport to measure affect, I find the same confusion which is found in the definitions of affect I have presented. The flood of measures purported to measure affect which has appeared in the past few years, the broad range of concerns and non-specific focus, and the psychometric inadequacies of the instruments suggest to me that: 1) We are not close to an agreement about what affect is or what to call it, 2) What we are trying to measure is so unclear to us that we cannot develop instruments with acceptable psychometric qualities, and 3) We probably will continue to enlarge our store of instruments which purport to measure affect until we can define what we mean by affect. I have concluded that unless we can achieve a better concept of affect, we will never be able to deal with it in our classrooms or in our research.

How were we trapped into the dualism which sometimes causes affect to be seen as an experiential problem and at other times as a physiological problem? In the time available I cannot fully answer the question but I will introduce it.

One thing we know is that the problem has been with us for a long time. It existed before Descartes postulated his mind-body dichotomy from which the experiential-physiological affect problem ensued. Had Descartes not divided us up into mind and body, I suppose someone else would have done so since long before Descartes we had accepted the philosophy of Socrates who rejected emotionality to free the rational mind for understanding and that of Aristotle whose logic posited a subject-object duality in all knowing. Were it not for the Cartesian dichotomy, we might think of the person we are trying to educate and not of the "mind" we are trying to teach or of the body which brings the "mind" to school.

Let me phrase the problem in another way. The Cartesian dichotomy has led us to an overconcern for bodily processes and physical manifestations of psychological processes since we cannot objectively study "mind." This exaggerated concern has, in recent years, taken the form of exclusive emphasis by the behaviorists on overt, observable, and measureable behavior. This over-concern with "behavior" has led to further lack of concern for what is happening to learners as people. In the past, there have been isolated voices urging us to do something about the whole person and to concern ourselves with what we are doing to the person of the learner. But these voices have been ignored for many reasons, the most recent of which is the conviction of many educators, which has come to education from psychology, that the only job of education is behavioral change.

Today we can no longer ignore affect. Federal legislation insists that we evaluate our programs at least in part on the basis of what is happening to learners' self concepts and values, among other things. And learners are showing us that we can no longer afford to ignore them. A few years ago they threatened to burn down our buildings. Today they have become disruptive, aggressive, hostile, and violent. In many ways, they have told us what we can do with our teaching, which many of them see as unnecessary, unimportant, irrelevant, and dehumanizing.

A second aspect of the problem which results from emphasizing the physical bases of emotion is that it diverts us from our basic concerns as educators and facilitators of change. If affect is emotion, then affect is a physical problem and not an educational one. To see affect as emotion is to ignore the affective experience and to concentrate on what seems to be its physical expressions. Let me phrase this in another manner. When I am angry or afraid, my experiential concern is not with the flow of blood to my large muscles, the dilation of my pupils, the increase of adrenalin in my blood stream, or other physical changes which, taken as a whole, are defined as emotion by most psychologists. My concern is with the experience of the emotion and this experience affects me as a whole. Thus, if the experience is sufficiently intense, it also causes changes in my physical being. It is interesting to note that when we look at emotion as a physiological process, not only do we ignore the experience of the emotion, we can't even determine what emotion is being experienced. At a physical level, fear and anger are surprisingly alike and love doesn't even exist.

We might ask why educators have chosen a direction that ignores what they call affect and which causes them to ignore some of their central concerns. There are several reasons. Let me share five of them with you.

1) One reason educators have avoided attempts to alter affect is the assumption that the inner world of a person is too private to enter; that it is a violation of human rights for schools to do anything about such things as attitudes, beliefs, values, and the like. This concern is based on the belief that the only way schools can assist the development of socially acceptable attitudes, values, and self concepts is to make students believe and value what we believe they should, and we are reluctant to move consciously into 1984 and political indoctrination. If this is the only alternative, then I agree; but I do not believe that indoctrination is the only alternative.

2) The lack of concern for affect in our classrooms has resulted also from the belief of many educators that they should be concerned only with behavior, since it is the only thing that really matters. Such a belief underlies most learning theories and theories of instruction. It is basic to many recent educational practices such as programmed instruction, computer assisted instruction, behavioral objectives, and objectives-referenced tests. In part, our overconcern for behavior has resulted from the efforts of psychology and education to become scientific. It is commonly held that the only way to develop psychology as a science is to concentrate on the study of behavior since only behavior can be observed and measured directly. Education has followed in psychology's footsteps.

3) The behavioral emphasis in psychology and education has resulted, also, from the matter of definition I raised earlier. Behavioral orientations define affect as an unfit subject for scientific study. Since science can deal only with rationality and, by agreement from the time of Socrates, affect and emotion are irrational, it follows that science cannot deal with affect. Thus, psychology very early gave up its concerns for soul and mind and divorced itself from philosophy. Psychology turned from a study of mind not because it was unimportant but because of the belief that the "inner man" could not be studied scientifically. The result in psychology and later in education was that inner experience was ignored.

4) A further reason for ignoring affect in the goals of the schools is the incorrect assumption that there is a direct

relationship between cognition and behavior. Schools have a singular concern for cognitive learning. Even when educators have sought goals for students such as creativity, critical thinking, worthy use of leisure time, worthy home membership, health, economic sufficiency, and others, it is usually assumed that the way to achieve these is through increased cognitive understandings. It is convenient to believe that we can reach the goal of a fuller life by increasing cognition, and this belief answers the question of how schools can help students reach the goal. Behaviorism helps to avoid the problem even more easily since it sees affect and other internal states either as problems which stand in the way of studying behavior or as non-existing or non-essential constructs, the latter position taken by B. F. Skinner. Skinner holds that learning is only the modification of already existing behaviors and that to change behavior, it is not necessary to know what is inside the "black box" called a subject if, indeed, there is anything there at all (Skinner, 1939, 1948, 1971).

5) In recent years, schools have been deterred from efforts to deal with affect because of the opinions of some parents and most critics who believe that social and other goals are diversions, dilutions of educational efforts, or frivolities and that affect is not the province of the school.

It is not only the public, the critics, and tradition which are responsible for the point of view that education should be concerned exclusively with cognition; educators also must accept some of the responsibility. How often have we heard the belief verbalized that, "concern for the affective lives of students is a nice goal but how are we supposed to deal with affect and still do everything else we are supposed to do?" That usually ends the discussion even though we may believe the question more appropriately should be, "How can we do what we are supposed to do for boys and girls *unless* we are concerned for the totality of their lives including their affective lives?"

In spite of beliefs that affect is not the province of the schools, many symptoms tell us that all is not well with education, and these are forcing us to re-examine the role of the inner experience of a person in his education. Please note that I say forced; since if the schools were functioning smoothly, I think they would continue with their exaggerated interest in cognition and the exclusion of concerns for learners as people. Among those things which are causing us to give attention to affectivity are these: 1) Students are not behaving as we think they should behave. They are difficult to control, they challenge our authority and our persons, and they are destructive and violent. 2) Many people are convinced that students are not learning what they should learn. 3) It is becoming increasingly difficult to coerce, manipulate, charm, seduce, or otherwise motivate students into becoming the kind of people we think they should be or into learning the things we think they should learn.

These problems have been developing for a long time and they are paralleled by other problems. Let me cite some additional ones from my own work and try to put them in perspective (Bills, 1975).

To help children learn the cognitions which we believe they should learn as they progress in school, educators provide children fewer opportunities to make decisions for themselves and ignore them more and more in the making of decisions. The deteriorating human relationships which result destroy the helpful qualities of the teachers. Teachers are rated by students as low in their level of regard for them and in their congruence or personal honesty with them. Teachers are rated even lower on their ability to understand students as students understand themselves and on their regard for students as people of worth unless the students' behavior conforms to the teachers' desires. As a consequence, the longer students remain in school the more negative become their attitudes toward school, toward themselves, and toward other people. These developing negative attitudes are paralleled by the increasingly more negative attitudes of parents toward the schools (Bills, 1975).

School is an especially bad place to be if you are a boy because teachers will see you as less acceptable than if you were a girl. They will have greater difficulty in understanding you as you understand yourself, and they will be more apprehensive about your potentially disruptive behavior. Sometimes you are better off if you are black rather than white; your chances of being ignored are greater. The middle-class culture of the average teacher, who is also female, makes it difficult for the teacher to understand and to accept black or male students. Girls are accepted only if they conform to the teachers' ideals.[1]

Unconscious biases assure that children who hold values different from those of their teachers will receive lower teacher-assigned marks even though their achievement levels warrant higher ones (Finch, Finch, and Bills, Note 2). Looked at in terms of freedom to exercise choice and to control one's future, schools must be judged as oppressive (Frymier, Bills, Russell, and Finch, 1975a, 1975b). With relationships in the schools such as these, is it any wonder that students, and particularly black students, at the junior high level are rejecting the values of their teachers (Bills, Note 1)?

Such symptoms and problems are causing educators to be concerned for humanizing the curriculum, for helping children relate more effectively to each other and to their teachers through such activities as values clarification, sensitivity training, individual and group counseling, magic circle groups, and others, for movements toward the teaching of values, for attempts to educate character, and for other things designed to overcome the problems.

The increasing complexity of human problems which face education have caused a corresponding increase in our efforts to solve the problems. Many of our efforts, though, which supposedly are intended to express concern for children as people, are efforts to "adjust" children to learning environments which exacerbate already existing problems. These efforts often create new problems and most have been directed toward discovering new ways of getting children to learn more information. As a result, the attempted solutions often become part of the problem.

The search for solutions to the problems which face the schools has led us into the area of affect, and here many unwelcome surprises await us. The search is a painful one. The confusion which is present in psychological theory about the nature of affect is found also in measures of affect.

Let me be more specific and discuss four compendia of what can be called affective instruments in a somewhat superficial manner to illustrate both the confusion which exists and my earlier point about the different emphases in education and traditional psychology. That point holds that people who are interested in facilitating change in other people see affect as a state related to cognition or inner experience whereas traditional psychologists see affect as a physiological phenomenon. Although only one of the four compendia is classed in the affect area by its author, I believe all four are in the area as defined by current educational use. It is no accident that the three compendia which do not claim to relate to affect have been developed by authors outside the field of education. Educators are the principal users of the term affect to describe a broad area of experiential concerns.

---

[1] Unfortunately, references to the data upon which this and preceding statements in this paragraph are based are not available since they were collected in confidential school surveys.

The first reference is to Buros' *Mental Measurements Yearbook* (1972). As the name implies, the reference is a cumulative index which appears aperiodically and which describes educational and psychological measuring instruments under the broad term of mental measurement.

For all of its sophistication, *The Mental Measurements Yearbook* has not solved the problem of classification or definition. If you are to use it, you almost need the title of the instruments you wish to review before you can locate them since the reference is far more useable starting with the title of an instrument than by starting with a problem and asking, "What instruments are available in the area?"

Buros, which is usually so helpful to people with measurement problems, is of little assistance in the area of self concept, attitudes, self esteem, values, morality, and character education. Of the 1,155 references in the seventh edition (1972), fewer than 25 are even remotely related to the area with which I believe we are concerned and most of these fall into categories of personality, interest, attitude, and self concept. One reason Buros does not give a long list of affective instruments is that such instruments have not been published commercially and *The Mental Measurements Yearbook* includes references only to published instruments.

The problem of classification of instruments and the contrasting concerns of educators and psychologists can be illustrated by reference to a widely cited inventory of "affective" instruments edited by Beatty and published by the Association for Supervision and Curriculum Development (1969). In it, Beatty has a chapter on "Emotions: the Missing Link in Education" (pp. 74-88). That chapter is concerned with self concept theory and not with emotion in the traditional psychological sense. Beatty believes that feelings and emotion are understandable only in relation to self concept. He states, " . . . feelings arise as a result of comparison between . . . incoming data and the self concept" (p. 81), and he differentiates between feelings and emotions primarily on the basis of intensity. He also sees feelings and emotions as different from each other but he believes that feelings become emotions when they become strong enough. Thus, there is a relationship between feeling and emotion, but feeling is the result of evaluation of self concept interactions whereas emotion is a physiological reaction. Beatty does not make it clear whether emotions and feelings differ only in degree but not in character. He does make it clear that feelings ary educational problems in the larger sense of the term "educational."

The inventory portion of Beatty's booklet illustrates the point of the confusion which has resulted from a lack of a clear definition of affect and suggests that Beatty's views about the nature of affect have not affected his list of affective instruments, since the inventory includes far more than self concept measures. This conclusion is further reinforced by the fact that although Beatty defines affect as already stated, that is, "a result of comparison between . . . incoming data and the self," he operationally defines it by the categories under which instruments are inventoried in the compendium. These include: attitude scales, creativity, interaction, miscellaneous, motivation, personality, readiness, and self concept. Thus, most assessment instruments which do not purport to measure achievement, aptitude, intelligence, or interests are included as measures of affectivity and affectivity becomes something other than what he defined it to be.

Robinson, Rusk, and Head (1968) speak only of attitudes, and these are described as political attitudes to which are appended two sub-groups, measures of occupational attitudes and occupational characteristics and measures of social psychological attitudes. You should not let the titles of the various sections of Appendix B, *Measures of Social Psychological Attitudes* (Robinson & Shaver, 1975), mislead you into believing that it contains descriptions of instruments which educators would expect to find there. Appendix B contains references and descriptions for most of the instruments which are of interest to educators when they refer to affect. It includes Measures of Self-Esteem and Related Constructs, Alienation and Anomia, Political Attitudes, Values, General Attitudes toward People, Religious Attitudes, and Methodological Scales. Methodological Scales include measures of social desirability and various social attitude scales. In case you are interested, "anomia" is defined by Srole, who authored one of the instruments, "as an individual's generalized, pervasive sense of social malintegration or 'self-to-others alienation' " (Robinson & Shaver, 1975, p. 172).

The Robinson and Shaver publication illustrates the fact that psychologists do not define affect in the same ways educators do. Although they include most of the constructs educators consider to be in the affective realm, these are not defined as affect measures but as attitude measures. Their method of grouping is not entirely satisfactory, however, since it uses constructs such as attitudes and values which are as poorly defined as affect. But I do not believe their groups are as likely to be as misdirecting to educators as the term affect. The reader is cautioned not to make final judgments regarding the instruments which are listed in Robinson and Shaver without reviewing the literature. This statement can be applied in varying degrees to each of the compendia to which I will refer.

In my particular field, which is self-concept theory, the most useful compilation is the one authored by Wylie (1974). Her book is well-written, clear in its definitions and standards, and it delimits the area so that it can be covered both intensively and extensively. Most importantly, Wylie assists the reader by specifying her criteria, definitions, and standards and then discusses the instruments in terms of these. She leaves no doubt in the reader's mind about her opinion of each of the instruments. Furthermore, at every possible point she attempts to point out the need for further research and makes numerous suggestions for improvement in the area.

Wylie's book is remarkable when it is considered that she does the work herself without the aid of a large staff. As a result, her information and conclusions are more reliable than those of any other list with which I am acquainted, although there is little doubt that error is contained even in this volume because of her need to depend to some degree on descriptions and data provided by authors.

Wylie does not encounter the problem of definition of affect, since her concern is with the self concept, which falls in the area of affect primarily through the thinking and practice of educators. A partial list of instruments which she reviews gives an insight into her thinking and includes measures of the phenomenal self, such as the Butler and Haigh SIO Q-Sort, the Index of Adjustment and Values, the Piers-Harris, Coopersmith's Self-Esteem Inventory, and Gough's Adjective Check List, and non-phenomenal self measures such as the Who Am I, Twenty Sentences Test, the Thematic Apperception Test, Rorschach Test, Draw-a-Person Test, and Sentence Completion Test. Although her concern is for the phenomenal and non-phenomenal self concepts, educators are interested principally in the phenomenal self which is concerned with the self as an object in the perceptual field and which is most often called self concept, rather than with self as agent of action which is the non-phenomenal self.

These four compendia give some idea of the inventory resources available to the researcher. They also reinforce the conclusion that there is no common agreement about what affect is, or of the types of instruments which should be used to measure it. The contents of the four compendia lead to several conclusions: 1) there is no common agreement about what to include in such inventories, probably because no broad and adequately defined concept is available to guide the selection of instruments, 2) most of the available instruments are of doubtful reliability and unknown validity, 3) most of the instruments in the area have been developed for use in one study or a series of studies and have then disappeared from use, and 4) even though an instrument is well-known and widely used, such as the Tennessee Self Concept Scale, the Piers-Harris, or Coopersmith's Self-Esteem Inventory, there may be little evidence to support its use. If you doubt the validity of this statement, I invite you to read Wylie's comments regarding the scales and especially the Tennessee Self Concept Scale.

From the thoughts I have shared with you to this point, I conclude that the answers to many of the problems in the affect

21

area are not to be found by plowing ahead in the same rut we have been travelling. What we need is a new frame of reference for understanding people and for viewing educational goals. The old frames of reference make allowance for affect only by tacking constructs on to already completed houses. As I said earlier, the reactions of busy teachers to such an approach is, "How can we do this and still do all of the other things we *must* do?" Permit me to return to Descartes in an effort to resolve the problems I have raised and to describe what I believe is a more productive point of view.

The mind-body dichotomy has affected our thinking in many ways. One of its more obvious manifestations, as far as education is concerned, is the recent attempt to develop taxonomies of educational objectives in the psychomotor, cognitive, and affective domains (Ragsdale, 1950; Bloom, 1956; Krathwohl, Bloom, and Masia, 1964). The taxonomy of objectives in the affective domain refers to these objectives as "emotional learning" objectives and classifies them into five categories which include Receiving, Responding, Valuing, Organization, and Characterization by a Value or Value Complex.

The five categories of the affective domain do not seem to me to show clearly that the taxonomy separates the cognitive and the affective domains, since cognition appears to be a part of some or all of them. Additionally, the taxonomy does not relate affect to emotion as promised by the term "emotional learning."

The dilemma is this: if you start with the mind-body dichotomy, no satisfactory scheme is available for logically relating affect to either part of the dichotomy, and affect sometimes appears as a "mind" concept and sometimes as a "body" concept. Educators view affect as an aspect of "mind" even though they use the terms affect and emotion interchangeably. On the other hand, psychologists who study emotion view affect as a "body" process. More precisely, those people who are concerned with helping relationships and with facilitating change in people, view affect as an aspect of a cognitive process whereas those who are interested in the scientific study of emotion view affect, or emotion, as a physiological process. Except for the term affect, we have no means of relating such concepts as values, attitudes, and self concept to each other even though we intuitively believe such a relationship exists. The term affect appears inadequate for the task.

A more satisfactory means of resolving the problem may be to begin by forgetting about the mind-body dichotomy and to view the person from a more wholistic point of view, such as is done by perceptual psychology. According to perceptual psychology, as a person interacts with other people, situations, and objects, certain aspects of the experience are differentiated as meaningful in either a positive or a negative sense. These differentiations and their meanings are called perceptions. Some perceptions are further differentiated as being a part of self while others are perceived as not belonging to the self. Perceptions and their meanings are organized into what is called the perceptual field which has a fluid and continuously changing character.

Central to the perceptual field is the self concept, which serves as a screen or sieve through which new perceptions are admitted to the perceptual field. Perceptionists (Snygg and Combs, 1949; Bills, 1956; Beatty, 1969) see behavior as a function of this perceptual field. Thus, all behavior is relevant to and determined by the structure of the perceptual field at the instant of action (Snygg and Combs, 1949).

For the person, his organized perceptions are reality and they are the only reality he can know. They have resulted from his interpretations of his interactions and especially those which involve his self concept. Interactions which are perceived as strengthening or confirmatory to the self are accompanied by positive affective states such as pleasure, joy, love, and others. Interactions which are perceived as potentially destructive or debasing to self or to the organism itself are accompanied by negative affective states and if the threat to the self is sufficiently great, the organism alerts itself for "fight or flight" (Cannon, 1932). The preparation for fight or flight involves physiological changes called emotions. Affect is the feeling state which results from the person's awareness of the state of his self concept.

The perceptual organization of a person includes all of his perceptions or beliefs about the nature of reality. Some of these beliefs are about what the nature of the physical universe is like, some of them are about what is important, others are beliefs about how we should be prepared to experience people and events, and some are beliefs about who he is and what he is like. These beliefs are called facts, values, attitudes, interests, prejudices, self concept, and the like. They are *all* beliefs and for the individual they are reality. They are all cognitive-perceptual in nature; they are the results of highly selective processes and definitions, and they have all been learned. In no way are they affective. Affect develops as a consequence of the interactions of the perceptual field into which they are organized with people, situations, and objects, and thus, affect is a relational result.

For our discussion the most important points are these. First, the individual is more profitably viewed as a total entity than as a combination of parts. Second, when viewed from this vantage point, affect is seen neither as an aspect of mind or body, nor as a type of cognition or behavior, but as the evaluative reactions of the organism during, and as a result of, interactions. Thus, affect is not a thing to be ignored, taught, or changed; it is the individual's perceptions or awareness of the state of his organism as it interacts.

A third point is that a perceptual point of view provides an understanding of why we have been unable to find common threads in what we operationally have defined as affective variables. From a perceptual point of view, the various things which we have called affective variables are not affective but are cognitive-perceptual beliefs. All cognitions are beliefs and they can be grouped in various meaningful ways. Some are beliefs about the nature of reality, while others are beliefs about what is important, about how we should feel about things and people, about what we are like, and so forth. All have been learned in that they have been differentiated in interaction. Even our beliefs about reality are learned and reality, as we believe it to be, exists principally through definition.

May I illustrate this last point through numbering systems? Before we can count, we must agree on the units we will use, the base of the system, which in the decimal system is 10, and whether or not the system will be finite or infinite, which means that when you reach a certain point you either begin over again or you continue on indefinitely. Arithmetic is based on an infinite, decimal system in which the units are equidistant from each other and are of equal size. But who is disturbed when we use a finite numbering system which has equal units and which has a base of 12, or when we use a finite numbering system with base 12 which has unequal intervals? The best example of my first illustration is a clock and of the second is a calendar. All three of these numbering systems -- our usual numbering system, the clock, and the calendar -- are acceptable and their acceptance lies in the definitions we have adopted.

My point is that all cognitions are beliefs and they can be grouped in various ways. One group is concerned with the so-called real world. The rest of our beliefs we have relegated to what we have called affect. In a traditional sense, cognitions relate to knowledge about the nature of the real world whereas those relating to what is supposedly more subjective have been classed as affect variables. The difference in the two groups is one of definition, and from a perceptual point of view, the groups differ only in terms of what they relate to. Thus affect, in the sense of the term as used by educators and others, is cognitive in the same sense that knowledge is cognitive.

To avoid possible confusion in what I will say, let me provide some definitions at this point. Our beliefs are perceptions in the sense that " . . . perception refers to any differentiation a person is capable of making in his perceptual field whether or not an objectively observable stimulus is present" (Combs, Richards, and Richards, 1976, pp. 16-17). Traditionally, the term "cognition" has been used to describe knowledge, understanding, or ideation, while in education, values, attitudes, self concept, and similar constructs have been called affect. The constructs of cognition and affect, when used in the sense I have just defined it, are both perceptual in nature. Thus, I subdivide perceptions into cognitive and non-cognitive types which frees the term affect for use as I have defined it. In my further comments, I will use cognitive, non-cognitive, and affect in a manner consistent with these definitions. I separate beliefs about reality from other beliefs and call them cognitive and non-cognitive only for purposes of discussion. They are all perceptual in nature; each affects the structure and content of the perceptual field, and through the person's perceptions, each affects behavior.

My fourth point in proposing that the perceptual point of view is a more effective way of viewing the human organism and behavior is my most important one. Viewed in this manner, affect should not be our most important concern if our purpose is to assist learning, development, creativity, and so forth. Our concern should be with the experiencing of the organism and with the availability of that experience in awareness and in undistorted fashion to the person as he interacts. It is at this point that our developing concerns for humanistic education, non-cognitive variables, group interaction, individual adjustment, values, and the like can be interrelated and brought into focus. Let me say how this works.

We know that if a person experiences another person as understanding, accepting, reliable, and trustworthy, then the ability of the person to admit into awareness any and all aspects of his experience is increased (Rogers, 1957, 1958). Threat reduces learning *and* the intellectual quality of behavior, as judged by its appropriateness and its ability to solve problems (Bills, 1969).

As you recognize, I have substituted one set of concepts for another and since I have, it is incumbent on me to show the relevance of the substitution for the problem of affective measurement in the sense that term is usually used. Unless I can show the relevance of my substitution, there is no benefit to be derived from this conceptual exercise.

Perceptual theory implies that our primary concern should be for those things which tell us something about the qualities of the relationships students experience and not for what we have called affect. This follows from what I have already said about threat and its influence on the structure of the perceptual field and through it, behavior. If we know something about the nature of the student's perceived environment in terms of potential threats or relationships which can reduce threat, then we can predict what will happen to the student (Rogers, 1959). Under reduced threat, a person lowers his defenses and begins to examine the meanings of his experience for him. His perceptual field broadens and he is able to admit to awareness more and more of his experiencing and is able to incorporate new experience. Consequently, his self concept alters to become more congruent with the meanings of this larger body of experience for his organism.

Under threat a person admits to awareness only selected aspects of his experience and even these may be distorted. As a result, he may come to view himself as inferior or superior to other people or he may view himself in other distorted ways. None of these perceptual states results in positive benefits for him. He not only is dissatisfied with himself and experiences painful interactions with other people, he is unable to learn as effectively and meaningfully as he might, and he is unable to use the totality of his experience to help him approach problems as intelligently as he might. Thus, the nature of the learning climate is a most significant consideration. If the learning climate increases personal threat, it increases problems and symptoms of problems. If it decreases personal threat, it enables the person to develop, to learn, and to become his experiencing. The nature of interpersonal relationships is central to the quality of the learning climate. If relationships are such that students perceive teachers as unconditionally positive in their regard for them, as understanding them as they understand themselves, and as congruent or honest, then the level of personal threat is low and students move toward becoming their experience. Thus, we would have two concerns: 1) what are the self concepts of our students, and 2) what is the nature of the psychological climate or the relationship climate we are providing students?

It is because of concern for the nature of the relationships which we provide students that I have sought to develop instruments such as the Locus of Responsibility Scale which attempts to answer the question of who is making the decisions in a classroom, the Relationship Inventory which measures the qualities students perceive in their relationships with teachers, the Feelings About School which is a global measure of students' beliefs about the qualities of their schools, the Parent Inventory which measures parents' perceptions of a school and their evaluations of it, the Teacher Problems Q-Sort which gives a measure of a teacher's openness to experience and ability to provide a helping relationship for students, and other measures which also might be classed as measures of the perceived qualities of learning climates (Bills, 1975).

It is also for this reason that I have sought to develop a series of self concept measures which I have called The Index of Adjustment and Values (Bills, 1975) and which give some idea of how a person perceives himself and other people. Measures of the self and measures of learning climates can be helpful in at least two ways.

I believe the most important of these is the determination of the quality of the learning climate we provide students. We can do far more to control the quality of the environment we offer children than we can to control their inner qualities or their behavior. Most psychotherapies demonstrate that if the climate in which a person is learning is positive, the self learnings also will be positive. Under such circumstances we know that people move toward a valuing process rather than the fixed and static values which characterize the middle-class climate of our schools. We know, too, that people under reduced threat are more likely to discover personal meanings in their learning and such meanings directly affect the intellectual level of their behavior. Another important aspect of this point takes us back to what I have already stated. Self perceptions develop in interaction and it is only in further interaction that we can change them. Thus, it becomes important to know more about the nature of our interactions with students, as seen by them, in order to improve the psychological climates we are affording them. I will say more about how we have attempted to do this later.

Probably, the most important use of measurement within perceptual theory is to assess our effects on students. For this reason, self concept instruments are particularly important. Please notice I am suggesting the use of these instruments to assess our effects and not to evaluate students. It is my opinion that most instruments, including measures of intelligence and achievement, should be used to assess our effects on students and not to assess learners. We should assume that learners have done as well as they can given their previous experiences and the quality of the experiences we are offering them. The evaluation, therefore, should be directed at the quality of our offerings while still remembering that students' learning is not entirely dependent on our effectiveness since it also is dependent on their past experiences and present perceptions.

The perceptual point of view can help to guide us in the selection of other non-cognitive instruments for understanding our effects on students. Some of these instruments measure attitudes, others measure values, and others measure interests.

I appreciate this opportunity to present my point of view and will welcome your comments and suggestions as an opportunity to further clarify my attempt to understand a very complex issue.

PROBLEMS IN THE MEASUREMENT AND USE OF AFFECTIVE DATA

At this point, I want to talk about problems in measurement and use of what usually is called affective data, but which I will call non-cognitive for the reasons I gave earlier. To do this, I have divided my comments into two sections. The first section is concerned with problems which arise when we seek to promote change in schools through the use of non-cognitive instruments. The second section is concerned with some of the formal problems of measurement in the area. I have chosen to emphasize the first section which deals with the use of non-cognitive measurement as an aid in improving schools because this approach helps schools while it provides important research data of a descriptive nature. I will say more about this later.

Earlier, I described a major problem of measurement in the non-cognitive area. That problem relates to both sections of my present topic. It is the problem of the lack of agreement about the nature of affect. Without an understanding of what affect is, we lack a basis for knowing what we should attempt to change, of what to measure, and of what instruments we should seek to develop.

Another problem which concerns people who are interested in non-cognitive measurement relates to the large scale use of measuring instruments because of difficulties in their application and subsequent processing. Closely related are the problems of when to use surveys or assessments to promote change and how to use their results. Another problem stems from the self report nature of most non-cognitive instruments and other inadequacies which are known to exist in instruments of this type. Additionally, there are legal problems which are related to self disclosure and the invasion of privacy.

I would like to begin my discussion of these and other problems by describing some of the things I have engaged in over the past several years. I believe that I can put the problems and my tentative solutions in context through this approach.

I have spent much of my time during the past eight years in developing a package of instruments in the non-cognitive area and in using the instruments to assist schools in their self improvement efforts. One of these instruments dates back to the early

1950's and the earliest forms of all were developed in the 1950's. I do not know the exact number of people we have tested with these instruments since 1969 nor the number of tests which have been administered, but a conservative estimate is that one or more of these non-cognitive instruments has been given to at least 100,000 people for a total of at least 250,000 separate tests. The purposes of the testing have been diverse and they indicate something about the extent of interest in the non-cognitive area as well as its variety.

In two large, metropolitan, county-wide school districts, the testing was part of a larger survey designed to help the central administration understand the problems of the schools and devise plans for their correction.

In another large, metropolitan, county-wide school district, the package was again a part of a larger survey but the problems which generated the survey were occasioned by desegregation of the schools. The school district had been given a series of integration orders by the courts over a period of several years and problems had resulted. Cross-district bussing and the pairing of schools was extensive and in one case a school had been designated to contain only grade eight. You can imagine the problems encountered by that particular school. Standards of behavior could not be passed on by older students to incoming students; teachers did not know the students as individual people; and there was a serious discipline problem.

Because of the confusion and because there was no commonly accepted direction for the further development of the system, the court decreed that no changes would be made in the school system until an unbiased agency had made recommendations for further change and development of the system. Our Bureau of Educational Services and Research entered the picture through a contract with the school board and completed its work when we reported back to the school board at a public meeting.

Another problem also came to us from a large, metropolitan, county-wide school system in its request for service. The system had built four large, open-space elementary schools and these had become the center of a heated controversy. Patrons were divided in their strong opinions about them. So our Bureau was asked for assistance. My problem was to study the attitudes of parents and students, the climate for learning, and the students' self concepts and achievement in the four schools and to compare these with results from four, matched, traditional schools in the system.

In another project, we tested the parents and children in 12 recently developed private schools. This effort was financed by a foundation which was concerned with the quality of these schools and with their impact on children. The schools participated for a variety of reasons such as desires for information to use in improvement efforts, desires of some headmasters to demonstrate their schools' superiority, and as a political device to convince parents that even though their children were in private schools, they did not lack access to the resources of a university. One headmaster wanted the data to prove to his board that he was right and they were wrong. I don't know if he proved his point and, if so, what happened to him.

Our more recent work has been through our General Assistance Center which is funded to assist schools with problems arising from sex and race discrimination. In this work our purpose is to provide a data base on which the schools can build a list of priorities for improvement.

One application was in a city school district which included 92 schools. Originally, the purpose of the survey was to provide information to the central offices for establishing priorities for school improvement. The purpose was changed, after a review of the data, to providing information for school by school improvement efforts, and this was later augmented by the findings of a study of violence in the schools which we conducted. The results of the violence study were interpreted in light of the survey findings and provided information to be used by the schools in combating violence.

The instruments have also been given in about 50 schools in addition to those already cited, principally for four purposes: 1) to standardize the instruments and to study their psychometric characteristics, 2) to provide information to schools for their use in improvement efforts, 3) to provide training opportunities for our graduate students, and 4) to collect data for research problems.

In these studies we have usually included the Parent Inventory, the Feelings About School, the Locus of Responsibility Scale, the Relationship Inventory, The Index of Adjustment and Values, and the Teacher Problems Q-Sort (Bills, 1975; Bills, Macagnoni, &

Elliott, 1964). In the larger school systems we used 15 to 20 per cent samples of students and parents and in all other schools we have used the entire school population. All instruments have optically scanable answer sheets. Our processing methods yield scores on each instrument and these scores can be produced by any of the demographic variables available which usually include sex and grade and, sometimes, race. Except for The Index of Adjustment and Values, the processing also produces copies of the instruments which show the number and percentage of students or parents choosing each response alternative.

From applications such as those I have described, we have learned a number of significant things. I would like to share seven of these with you.

1) My first generalization is this: Change results most frequently from a survey when it has been requested by a school system's central administration and the results are organized so that the remedial or improvement efforts which the administration can take are clearly emphasized. If the purpose of a survey is to provide information for the local schools to use in their improvement efforts, change is less likely to occur.

2) If a school principal requests a survey simply because he believes it is important, and if he does this without consulting and fully involving the faculty, it is difficult, and oftentimes impossible, to get the faculty to use the results. Under these conditions, the faculty most frequently defends itself against the necessity to change by challenging the validity of the results. I might add that there is no adequate defense against such a challenge. In our testing, the faculty administers the tests and they can cite many examples of conditions which invalidate the data. They obviously do not feel responsible for the invalid data. As they see it, they are not responsible since the collection of the data was not their decision. I do not believe it is possible to force change on a school.

An example of the difficulty in initiating change without going through a process in which involved people collect data which are meaningful to them comes from the testing we have done in the past which had as its primary purpose the instruction of our students. In such instances we reported the results of the survey to the school, we interpreted the results, and we made suggestions for improvement. The faculty listened politely and then ignored the findings and recommendations. This is not a criticism of the teachers. The purpose of the teachers was to assist us and our students. The teachers acted conscientiously in the collection of the data to provide a good laboratory experience for our students. Their purpose was not to use the survey for uncovering problems and possible solutions. So we can generalize that when an attempt is made to involve people in something which someone else believes is meaningful, but which is not perceived by them as meaningful, the results are usually inconsequential. Teachers more frequently try to improve their schools when principals help them to determine if there are problems and if they need additional data to understand the problems and to solve them.

3) For similar reasons, we have learned that the feedback we give teachers must be directly meaningful to them. If the information is personally meaningful, teachers are more likely to change themselves than if they do not perceive personal meaning in the data. Consequently, we give each teacher individual reports which tell how one class described the teacher on the Locus of Responsibility Scale and the Relationship Inventory (Bills, 1975). The reports consist of the seven scores from these two instruments printed out student by student along with the class average and the norm averages. Included in each report is a print-out of the test along with the number and percentage of students who selected each response alternative. Whenever possible, we make individual appointments with the teachers and go over the results to aid interpretation and to help the teachers understand how they are seen by their students. Teachers do not leave such conferences with the same perceptions as when they entered. Incidentally, the children are never asked to identify themselves.

It is a joy to report to a teacher who is providing students with opportunities for making decisions for themselves, assisting students in making these decisions, and providing them with a climate in which the students dare to try new things. Our most positive results may come from those teachers who are sincerely trying to help students and who have little information with which to evaluate their efforts. Such information seems to encourage these teachers to try even harder.

Teachers with inadequate relationships with their students tend to become defensive in these reporting sessions. When this happens, our best counseling skills are called for and sometimes we can help the teachers to lower their defenses and to recognize what they really have known all along but could not admit to themselves – that things are not going the way they want them to go. They then begin to explore how they can change.

We have had a few surprises while counseling what are obviously poor teachers. Some of these people are able to admit to themselves what they already had suspected – that they are ineffective and that they should not continue to teach. In the few cases in which I have seen this happen, the teachers accept the decision as their own and they leave teaching without recriminations. Their feeling seems to be, "Teaching is important but it is not a place where I can make an important contribution."

4) My fourth generalization is that if you are working within a framework designed to promote change, it is important that commitments for follow-up efforts be elicited before the work begins. If commitments are not made at that time, there is little hope that they will be made following the presentation of the facts. Teachers do not become excited about results since there is so much else to do to which they are already committed. Too, the results of the surveys often are such that teachers want to ignore, to deny, or to forget them. The teachers want to believe they are already doing their best, and many of them are, but the survey may present evidence that even their best efforts are not producing the results they desire. Their impulse is to question the validity of the survey, the instruments, the interpretations, or anything else which will dispose of the results as quickly as possible.

Unpleasant results must occur at least one-half of the time because one-half of the cases will be below the mean. This can be a serious drawback to the use of normative data and we are seeing the consequences every day in the press. Who can fail to be impressed by a statement such as, "Half of the children in the public schools are reading below grade level?" The immediate interpretation is that schools are failing and something must be done about it. The uninformed critic makes the statement and the concerned patron wants to act on it. But if you examine the nature of such data you discover that "grade level" is defined as the average score made by a child at a particular grade level, thus, on the average one-half of the children will be below grade level and one-half will be above it. Have you ever stopped to think that one-half of all practicing physicians were in the bottom halves of their graduating classes? Facetiously, I could say that it is no wonder that half of the children in our schools are below grade level, after all half of them are below average in intelligence. But try to explain the fallacy to teachers who have tried what they believe is their best, who are frustrated because their efforts have not reached all of the children to the degree the teachers desire, and who are tired and exhausted from their day's efforts and are already looking forward to cooking the evening's meal which they may not want to do.

There is another factor which affects teachers making commitments to do something about the results of a survey after the results are known. If a variety of instruments is used, it is likely that the results will be below average on at least one of them, although the results may be above average on the others. People take consolation in the above average measures which makes it easier for them to ignore the below average ones.

If you test enough schools, some of them will turn out to be average. But who within our striving culture wants to be average? Being average may be as bad as being below average and the same result occurs as far as commitment to action is concerned.

But suppose a school is above average. Will this cause a commitment? The results probably will be the same as for the other groups. If you are above average, you may be able to breathe a sigh of relief and to forget the whole thing. On the other hand, being above average may be a threat for an overly-concerned teacher who desires perfection. The results will be harmful to the children if they cause the teacher to strive even harder. Our greatest hope for success, though, is this above average group. These people are usually less defensive. Consequently, they may be willing to do even more, but this may not be what we want. We may find ourselves spending much time with schools which are already doing well.

My fourth point, then, is this: commitments for follow-up efforts and for establishing priorities for improvement must be made prior to a survey.

Why is there no greater willingness on the part of teachers to make commitments than I seem to have indicated? I have given several answers to this already but I wonder if there isn't an even more important one. In a recent speech, Ralph Nader (Note 3) pointed to an important characteristic of economic theory. According to Nader, all economic theories are production-centered. This is true even for such diverse theories as those of Adam Smith, of Keynesian economics, and of socialistic economic theories. They are production-centered and the gross national product or a similar index is the criterion of success. An increase in the GNP is considered good even if occasioned by such a thing as an increase in crime which requires the purchase of more sophisticated detection devices, communication devices, and computer control, all of which add to the GNP.

Here is another example. The Department of Commerce has said that within the United States almost 600,000 workers earn $7 billion annually in the production and use of fluorocarbons, and nearly 1.5 million more have jobs that depend indirectly on fuorocarbon products and their applications. Most of this effort produces aerosol sprays which save us from such undesirable things as underarm wetness and odor and from unruly hair and the "wet" look while possibly destroying the ozone layer in the atmosphere and threatening us with a serious increase in skin cancer. This industry rests on high-powered advertising, developed by so-called consumer research. It adds significantly to the GNP without adding significantly to human welfare or the quality of life, all the while adding to the pollution problem and possibly having long range effects on climate.

Let me return to Nader. He stated that if we eliminated crime, waste, and pollution, our economy would become stagnant. He argued for a consumer-oriented economic theory and his speech was a challenge to America's psychologists to aid in this effort. Frankly, I think he was talking to the wrong group. Psychologists who are interested in the market-place and in industry align themselves with advertising and with management. Consumer psychologists are concerned with questions such as what influences choice in a supermarket? Their answers to the question are more useful to producers and marketers than to consumers.

I believe there is a direct parallel between our production-centered economic theories and the way we view education. Evaluation of education is production-centered – what is our retention rate, what is the average achievement of our product, and what is the cost per unit? Such questions do not suggest that education is consumer-centered and concerned with students as individuals. In many ways education is production-centered and it is becoming even more production-centered. Education is being judged by the quality of its GNP, not for what has been done for boys and girls. Additional efforts to improve the non-cognitive lives of children may result in a reduction of the quality of our educational GNP as measured by norm-referenced tests of cognition which are what schools are being judged by. Even teachers complain about the quality of the educational products turned out by the children's previous teachers. As painful as it is for me, I must ask, "Is education really dedicated to the enhancement of the lives of children or is it dedicated to the enhancement of educational production?" Some of you may say the two alternatives are the same; my belief is that they are significantly different. So my point is this: economics is production-centered; is education any the less so? Do we have real concerns for our learners as people? If so, why do we ignore these concerns so frequently, especially in secondary and higher education? To what degree has education's reluctance to be concerned with the non-cognitive learnings of children stemmed from this same production orientation?

5) My next point can be stated briefly. In an improvement effort, commitment by teachers and administrators can be achieved easier from the use of a few well-chosen instruments than from the use of a large battery of instruments. The instrument or instruments should be selected primarily on the basis of problems which have been identified as important. When problems have been identified and instruments have been selected, a faculty has taken the first steps in commitment. In a moment, I will illustrate the progress made by one very excellent school which used only one instrument for its study.

6) My sixth point is about the anonymity or confidentiality we must afford respondents in non-cognitive surveys. The obvious reason is that when we intrude on the personal lives of people, we must expect many of them to distort their disclosures because of the threat of adverse criticism and desires to appear adequate. If we are measuring cognition, we can ask for names and other identifying information. The student will do his best to give acceptable answers to the questions. He may do the same thing on a non-cognitive test and attempt to give socially acceptable

answers. The factor of social desirability in non-cognitive tests may be an asset or it may be a liability. It may be an asset to have a measure of a person's need to give socially acceptable answers, but it is a liability if it interferes with accurate measurement of other variables. Anonymity reduces threat for most people and, thus, the need to distort.

There are other cogent reasons for avoiding identification, including the right of the individual to his privacy. Asking people for non-cognitive information when they are neither free to give nor to withhold permission is a violation of ethics.

From a practical point of view, there is another important reason for not requesting people to identify themselves when they respond to non-cognitive instruments. The immediate impulse of many teachers when they see a list of scores with names attached is to see how Mary or Jimmy or Samuel or Susy are doing. And this sort of examination of data destroys improvement efforts. Let me illustrate the point. Teachers' reactions to IQ scores frequently are, "Well, I would have expected Mary to score high," or "No wonder Jimmy is doing such poor work," or "It's obvious that Samuel is doing better than we might have expected," or "Susy could do the work if she wanted to." None of these alternatives suggests that change is necessary. The identification of students may impede our efforts.

I shudder when I think of the fact that some of our data might have appeared on permanent record cards had names been attached to them. What a travesty. No test data that I know of are so reliable and valid that they should be written indelibly into the record of a person. When we do so we bind the person to his past instead of releasing him to his future. The school's job is to reduce the ability of tests and other devices to predict, not to enhance their effectiveness.

So in my efforts I assiduously avoid the identification of students by name. We do ask teachers to identify themselves since we wish to report the results to them in individual conferences or through confidential reports. But we do this only after the teachers are convinced that the results will remain confidential and will be shared with no one in the school system. I am usually able to convince teachers that the identifying information will be kept in strictest confidence. I also acquaint them with the fact that the numbers assigned to them for use in the testing will be reassigned before the data analysis to provide further protection.

I should point out that the data on which teachers can be identified come from instruments which are completed by the students. Students' efforts to distort such data are probably in the direction of attempting to emphasize their own perceptions of the teacher and not to falsify the truth as they believe it to be.

7) My seventh, and final, point is that the characteristics of the survey instruments determine, in part, their usability in change efforts. I will illustrate this sub-point by reference to The Index of Adjustment and Values (Bills, 1975), which I have worked longest to perfect and about which Wylie (1974) says:

> It is evident that the IAV has been used by many researchers. Reliability is quite high. Evidence for convergent validity includes correlations with many different purported measures of self-regard – a wider range of such instruments than is the case for any other self-regard measure. Although the degree of convergent validity of any of the self-regard scores from the IAV is quite moderate, it is probably as good as for any extant instrument which purports to measure global self-regard with the use of numerous items; and the IAV is shorter and less cumbersome for . . . (subjects) and . . . (experimenters) than is true of a number of other intruments . . . (p. 165).

I cite this quotation not to brag but to counteract any false impression of the IAV which my next remarks may create.

The IAV is a useful tool for researchers. The insights it has given me are rewarding and have helped me organize my understandings into meaningful relationships. When used as a part of an assessment battery for the purpose of developing a program for effecting change, it is probably the least productive of the instruments in our package. It is concerned with things over which teachers have little direct control. How much value is it to a teacher to know that the average child in his class is lacking in self acceptance and sees other people as having greater worth, to give an example of what might be reported? What is the teacher going to do about this in a direct manner? Is the teacher to give the children pep talks to help them believe they are as good as other people? Such approaches are not usually productive. Have you, for instance, changed your opinion about Alabamians because of Mr. Wallace's insistence that, "Alabamians are just as good as other people?" It hasn't affected my opinion of myself nor my opinion of Alabamians.

The most useful instrument in the package, for purpose of helping schools change, may be the Feelings About School (Bills, 1975). The Feelings About School is a 50 item true-false instrument which asks children to respond to statements such as, "I like the way our teachers treat us," "Our teachers are interested in us," "My classes are boring," "In general, I like my teachers," "Some of the teachers act as if they want the students to feel embarrased," and "There are too many rules in this school."

The value of the Feelings About School stems, in part, from the manner in which it can be used in improvement efforts. As I said earlier, children's responses to our survey instruments are reported to the school by grade or sex, combinations of these, or any other demographic variable requested by the school. For each group, we report the average score, the norm, and the tabulated responses of the children. Our suggestion to the teachers is to take print-outs of the Feelings About School which are appropriate for the grade levels at which the teachers work and to ask the children to amplify their responses. The teachers are asked to listen intently and to avoid personal reactions. The teacher reads an item to the children and then tells them how they responded. The teacher then asks, "Do you think this is the way the children feel about the school?" Most of the time the children will agree with the statement but often they will say the situation is worse now than it was. The teacher carefully avoids asking the children why they feel as they do. "Why" questions are threatening; they require some kind of personal answer which you are expected to be able to defend. Furthermore, the "why" question may lead to a discussion of personalities, usually those of the teachers, and what the children say may not be at all complimentary. Instead, the teacher asks the simple question, "What do you think the school might do about it?" At this point the teacher takes obvious and copious notes and is careful to avoid comments such as, "You know you don't believe that," or "That's a ridiculous suggestion. Has someone else a better answer?" Instead, the teacher listens, attempts to understand what the children are saying, and records their responses. I usually suggest that the teachers get together with consultative help and look at the responses, asking themselves, "What is the overall picture of what the children are telling us?"

In the very excellent school to which I referred earlier, I was teaching an off-campus course and had frequent contacts with the faculty. We looked together at what the children had said and then elementary and secondary teacher groups attempted to generalize from the particular suggestions of the children. The results were astonishing in a number of ways. The children told the teachers almost the same things a professional educator might have told them about the quality of human relationships, the type of instructional practices, the opportunity of the children to have a voice in their own education, the feelings of respect and affection the children wanted teachers to have for them, and others. A second interesting feature was that the elementary children were far more insightful about what should be done to improve the school than were the secondary children. The secondary children had begun to adopt the point of view of their elders that the purpose of school is to cover text-books and to teach information and that the role of the teacher is to tell the children what they need to know and to assist them in learning it.

The important part of the story is what happened after this exercise. I asked how the school might use the results and the teachers began to respond as the children had. I helped them generalize, from their responses. As a consequence, faculty meetings were given over to an improvement process which continued for the next two years. I don't know how effective the changes were that the school initiated, but I do know that the teachers saw the children differently and had formed a partnership with them to explore ways to improve. No doubt the children saw the teachers differently than before.

Reliability and Validity

You may notice that to this point I have avoided the terms reliability and validity. I have done this for several reasons. The most important stems from what I tried to establish earlier. Affect arises as a consequence of the interactions of a person with another person, situations, or objects. It is within the context of relationships that affect develops and it will change only through further and improved relationships. So my emphasis has been on the use of non-cognitive instruments as means of establishing improved relationships in which change can occur.

But there is another reason for my approach. Relationships between educational researchers and local schools have changed rapidly in the past few years and they promise to change even more in the future. Traditionally, research in the schools has been the province of educational psychologists and other faculty members in professional education and was conducted for a variety of reasons, including desires to collect research data with which to test hypotheses and to construct theories, particularly theories of how people learn, to develop a body of information and conclusions which could be used to construct and modify school programs, and to collect information to further understandings of human development.

Today, the nature of the relationships between educational researchers and the schools has changed. The public schools are more sophisticated in problem definition and designing approaches to problem solution than ever before and they are improving continuously. The schools must be concerned with and attempt to solve problems of definition, project design and implementation, and evaluation of outcomes to satisfy their publics and to secure the federal funds which are needed to initiate innovative programs. Public school personnel no longer are content to cooperate with our requests for use of the schools as laboratories for the collection of data relevant to problems which are of concern solely to us in higher education. What they want is assistance in solving their own problems, many of which are research problems. For this reason public school people seek our help and I believe that those of us in institutions of public higher education owe it to the schools to help to the degree we can.

Many institutions have found the new relationship a profitable one. We have a waiting list of schools which have requested our help. We find ourselves working on exciting problems which have a degree of importance in the lives of people which we never achieved in the older relationship. The new relationship has not been a one-way street, though. It has given us an excellent laboratory for educating our educational research students. There is no dearth of opportunity for working with people to define problems, to construct approaches to solve the problems, to select or design the instruments needed for evaluation, to assist in the collection of data, to analyze the data with our specialized resources such as optical scanning devices, computer facilities, and computer programs, and to assist in the interpretation of the data and the use of the interpretations.

But there is another important reason why I have relegated "pure" non-cognitive research to a secondary position in this paper. That reason rests on the present state of our knowledge in the non-cognitive area as it pertains to definition and instrumentation. Let me quote again from Wylie (1974):

> Although progress has been made in the last decade, no investigator has satisfactorily conceptualized or coped with all the difficult measurement problems in the self-concept field. Quite a few have indicated that they make no claims for having tried to support the reliability or construct validity of their instruments, and they are content to 'let the reader beware,' as it were (p. 123).

The same thing can be said about other non-cognitive measures.

The research emphasis in the non-cognitive area at the present time probably should be on defining constructs, developing instruments to measure these constructs, and validating instruments as measures of the constructs rather than with efforts to predict differences in groups, to measure change resulting from the application of independent variables, exploring relationships of non-cognitive variables to other variables, and other similar problems. The research problems in the area should be directed toward a determination of the nature of constructs which can most profitably be conceptualized as central to our non-cognitive concerns, toward more adequate definitions of these contructs, and the construction and validation of instruments to measure them. Without such definitions and instruments, we can hope to add little to the area except more confusion. We will add to the solution of the problems if we bend our efforts and those of our students toward definition and instrumentation and away from the further proliferation of crudely constructed instruments to measure vaguely defined constructs. Most of these instruments are never used again, but if they survive, they frequently are used by uncritical researchers and result in further confusion.

The problem of test proliferation results, in part, from a failure to understand the nature of theory. Too often we view theory as a set of truths which have not been completely validated, rather than as a means of conceptualization which attempts to place known facts into meaningful relationships with each other for purposes of understanding and prediction of new facts and relationships. Progress results when we push a theory to its limits, for then we can see the inadequacies of the theory and construct a new one. More often than not, though, someone believes he has insight into truth. So he constructs an instrument to collect data with which to prove this truth. The instrument, if it is ever used again, may eventually appear in a compendium of instruments. Since the instrument does not rest on well-defined constructs, it becomes known by its title, and it is used subsequently for measuring constructs inferred from the title. This eventuates in more confusion.

Frequently, the solution of a school's problem seems to require the development of a special instrument and often the construction of a new instrument cannot be avoided. However, it can frequently be avoided by examining other means of data collection and a more thorough examination of the problem may result in a redefinition of the variables so they can be measured with instruments already available.

In discussing reliability and validity it is important to emphasize that the use of an instrument determines, in part, the emphasis which must be given to its reliability and validity. If we are using instruments in change processes, questions of reliability and validity are frequently of secondary importance since we can compare the results with people's perceptions of their experience. If, though, we are involved in more traditional research efforts or if we are to base recommendations for change on the use of data collected by formal instruments, then reliability and validity are more important concerns. Most instruments in the non-cognitive area fail to pass the reliability and validity tests. True, there are some instruments which have reasonably high coefficients of consistency and stability, however, the reliabilities of many non-cognitive instruments are not known, and the reliabilities of most are unacceptable.

I have given one example of how use determines the importance of reliability. This was the means by which the school that used the Feelings About School used the data to change their instructional programs. The Feelings About School has excellent reliability but the reliability of the instrument probably is of secondary concern when it is used in the manner I described.

Here is another example of how use determines the importance of reliability. If an instrument is to be used for making judgments about a person which will significantly affect his future, then none of the non-cognitive instruments which are available pass the test in my opinion. If, though, we are using an instrument in a controlled experiment, we may accept one with less than optimum reliability although we must recognize that the lack of reliability always contributes to the error variance and many studies have failed to significantly alter educational thought because they failed to establish predicted relationships even though the failure may have been due to a lack of reliable instrumentation.

Although reliability is a necessary condition for the use of most instruments in research work, it is not a sufficient condition to justify their use. Of greater importance is the question of validity and in the non-cognitive area, as in the cognitive area, this is difficult to achieve. The primary reason for this is the lack of criteria against which tests can be validated. Before moving into the question of validity, let me clarify a statement which I just made.

I said that reliability is a necessary condition for the acceptance of most instruments. The word "most" in the statement relates to the fact that it is not possible to establish reliability for many useful instruments. For example, the reliability of a questionnaire usually cannot be established. On a questionnaire which seeks to gather information, why should we be able to predict the answer to one item from the answers to other items?

Other examples of instruments for which reliability coefficients are unobtainable or meaningless are in Q-technology. This widely used technique requires the construction of a Q-sort and its analysis by means such as intra-person correlation and factor analysis of the resulting correlation matrix. Q-technology is the inverse of R or Pearson product-moment correlation. In an R study, a group or population of people is given a number of tests after which the tests are correlated with each other and this inter-person matrix is then factor analyzed to see the factors which are common and unique to the tests. In Q, a population of tests is given to a number of people or to one person who is asked to describe himself in a variety of ways such as how I used to be, how I am now, what I would like to be, how I am seen by my mother, etc. Following this the various aspects of the person are correlated with each other and then factor analyzed. In this case each item of the Q-sort is a test. Since an item is a test, its internal consistency cannot be measured.

To me, Q-technology is the most useful technique available for quantifying descriptive data and for solving some of the problems I have discussed. It is an excellent means of generating hypotheses and has been used in a wide variety of ways such as studying changes in patients during psychotherapy (Rogers and Dymond, 1954), differences and similarities in the types of relationships experienced and inexperienced therapists of different persuasions attempt to effect with their clients (Fiedler, 1950, 1953), role concepts of university professors (Bills, 1975), the purposes of a college or university (Bills, 1970), and the openness of teachers to experience (Bills, Macagnoni, and Elliott, 1964).

What I have said is that reliability is usually an important feature of a measurement instrument although not always measurable and sometimes not even applicable. I have also said that the reliabilities of non-cognitive instruments vary from unknown, through unacceptable or barely acceptable, to a few with acceptable reliabilities.

So much for reliability. Let's take a quick look at validity and introduce it by noting that Kerlinger (1973) says that, "The subject of validity is complex, controversial, and peculiarly important in behavioral research. Here perhaps more than anywhere else, the nature of reality is questioned" (p. 456). It is impossible to address the validity problem without getting into philosophy. Inquiries into construct validity, in particular, are inquiries into the nature and meaning of variables.

Actually, I have already discussed what is probably the most important aspect of the validity question. Of the various types of validity the most important in the non-cognitive area is construct validity since it is in the area of construct validity that psychometry and theory meet. Whereas construct validity is most difficult to establish, other types present simpler problems. For example, it is relatively easy to determine if an instrument can predict with a better than chance degree of accuracy. But even if it does, this does not prove that it is a measure of what it purports to measure – in other words, that it has construct validity. There was a time when we were not concerned for the reasons behind predictive or concurrent validity; if an instrument predicted what we wanted it to predict, then it was acceptable. Ability to make predictions from unknown bases still characterize industrial psychology and motivation research. The Minnesota Multiphasic Personality Inventory rests its validity on its ability to predict the nosologic groups that hospitalized patients will fall into. Predictive validity may have value but it adds little to basic understanding and it is not helpful when we desire to determine if applications of newly derived educational methodologies positively influence the non-cognitive lives of students.

Convergence and discriminability have been suggested as means of establishing construct validity but I cannot see that they solve the basic problem either. If two instruments which purport to be measures of the same construct agree with each other while disagreeing with other instruments which purport to measure other constructs, we know something about the instruments but we do not know if any of them measure the construct.

Part of the confusion about which variables in the non-cognitive area are central, necessary, productive, and non-duplicative results from inadequate knowledge regarding the construct validity of non-cognitive measures and construct validity poses what may be an unsolvable problem. Since non-cognitive variables are constructs or concepts, it is doubtful that instruments with construct validity can be developed to measure them. The reason is simple. A construct is a means of conceptualization. It is not a thing, a truth, or an object; it is an idea or a miniature theory. To establish construct validity would mean that an idea or a concept is a reality and can be proved to exist. This has been a basic problem in the measurement of intelligence. Since we do not know what intelligence is, we cannot tell if a test measures it. It is possible to validate a test as a measure of intelligence only through devious and sometimes questionable means which establish predictive validity but not construct validity. Thus, most developers of intelligence tests shy from claims that their tests measure intelligence and they say they measure things such as IQ, mental maturity, or scholastic aptitude. A variety of imaginative means have been used in attempts to establish the construct validity of non-cognitive measures. Most of these involve such complicated logic that the researcher may not be certain of the type of validity which is being tested or if the test establishes any evidence of validity. The best we can hope for in validation studies is support for the construct as a useful means of conceptualization, but this supports only the usefulness of the construct, not its validity.

But even though we may not be able to solve the construct validity problem, it is possible to determine to what degree a test measures only one factor and to determine to what degree two or more tests overlap in what they measure. This can be done by means of factor analysis. Factor analysis can be used to develop instruments which are factor pure and to eliminate overlapping variables which are called by different names. Kerlinger (1973) claims that factor analysis can be used to develop instruments with construct validity although Cronbach (1971) disagrees. I also disagree for the reasons I have given. We would do well to help our doctoral students concentrate in this area.

Let me summarize. Surveys of instruments in the non-cognitive area show that some have unknown reliability, many have poor reliability, and a few have reasonably high reliability. These same surveys show also that little is known about the validity of non-cognitive instruments, that content validity is all that is claimed for most non-cognitive instruments, and that practically nothing is known about their construct validity.

It is my belief that we would do well to involve ourselves more with the first mode of operation which I have described in this paper – that is, in attempts to facilitate change in schools – than in what might be considered as more basic research until such time as better instruments are available. If educators want to get into the basic areas, let them aid in the solution of the problem by helping to develop more adequate constructs, definitions, and instruments. In the meantime, the descriptive research results which become available to us from efforts to effect change such as I have described are fruitful sources of hypotheses for further study. Such hypotheses can lead us toward better definitions and toward better instrumentation. Science usually starts with description and the science of the non-cognitive aspects of human experience could profit from more descriptive efforts.

REFERENCE NOTES

1. Bills, R. E. *Racial differences in perceptions of self and others.* Paper presented at the annual meeting of the Mid-South Educational Research Association, Jackson, Mississippi, November 1975.

2. Finch, J. D., Finch, C. B., and Bills, R. E. *Values and school achievement.* Paper presented at annual meeting of the Mid-South Educational Research Association, Memphis, Tennessee, November 1972.

3. Nader, R. N. Bringing psychology into the consumer movement. Speech to 84th Annual Convention, American Psychological Association, Washington, D.C., September 3, 1976.

REFERENCES

Beatty, W. F. (Ed.), *Improving assessment & an inventory of measures of affective behavior*. Washington, DC: Association for Supervision and Curriculum Development, 1969.

Beigel, H. G. *Dictionary of psychology and related fields*. New York: Frederick Ungar Publishing Co., 1971.

Bills, R. E. *About people and teaching*. Lexington, Ky.: Bureau of School Service, College of Education, University of Kentucky. *28* (2), 1955.

Bills, R. E. Love me to love thee. *Theory Into Practice*. 1969, *8* (2), 79-85.

Bills, R. E. *The study of institutional purpose by Q-technology*. University: The University of Alabama, Bureau of Educational Service and Research, *1* (1), 1970 (Nov.).

Bills, R. E. *Improvement of instruction through teacher description*. University: The University of Alabama Venture Fund, 1975a.

Bills, R. E. *A system for assessing affectivity*. University: The University of Alabama Press, 1975b.

Bills, R. E., Macagnoni, V. M., and Elliott, R. J. *Student teacher personality change as a function of the personalities of supervising and cooperating teachers*. Final Report on Project S-020. U.S. Office of Education, Cooperative Research Program, 1964. ERIC reference: ED 003441.

Bloom, B. S. (Ed.) *Taxonomy of educational objectives. Handbook: I. Cognitive domain*. New York: David McKay, 1956.

Buros, O. K. (Ed.), *The mental measurements yearbook. Vol. 1 and 2*. Highland Park, NJ: Gryphon Press, 1972.

Cannon, W. B. *The wisdom of the body*. New York: Norton, 1932.

Combs, A. W., Richards, A. C., and Richards, F. *Perceptual psychology*. New York: Harper, 1976.

Cronbach, L. Test validation. In R. Thorndike (Ed.), *Educational measurement*, 2nd ed. Washington, DC: American Council on Education, 1971.

Duncker, K. and Watt, D. G. *A German-English dictionary of psychological and psychoanalytical terms*. Ann Arbor: Edwards Brothers, 1938.

English, H. B. and English, A. C. *A comprehensive dictionary of psychological and psychoanalytical terms*. New York: Longsman, Green, 1958.

Eysenck, H. J., Arnold, W., and Meili, R. *Encyclopedia of psychology. Vol. 1, A to F*. New York: Herder and Herder, 1972.

Fiedler, F. E. A comparison of therapeutic relationships in psychoanalytic, non-directive, and Adlerian therapy. *Journal of Consulting Psychology*, 1950, *14*, 430-445.

Fiedler, F. E. Quantitative studies on the role of therapists' feelings toward their patients. In O. H. Mowrer (ed.), *Psychotherapy: theoretical research*. New York: Ronald Press, 1953.

Frymier, J. R., Bills, R. E., Russell, J., and Finch, C. B. A study of oppressive practices in schools. *Curriculum Theory Network*, 1975a, *4* (4), 309-313.

Frymier, J. R., Bills, R. E., Russell, J., and Finch, C. B. Are the schools oppressive? A preliminary report. *Educational Leadership*, 1975b, *32* (8), 531-36.

Kerlinger, F. N. *Foundations of behavioral research*, 2d ed. New York: Holt, Rinehart, and Winston, 1973.

Krathwohl, D. R., Bloom, B. S., and Masia, B. B. *A taxonomy of educational objectives. Handbook: II. affective domain*. New York: David McKay, 1964.

Robinson, J. P., Rusk, J. G., and Head, K. B. *Measures of political attitudes*. Ann Arbor: University of Michigan, Institute for Social Research, 1968.

Robinson, J. P. and Shaver, P. R. *Measurement of social psychological attitudes*. Ann Arbor: University of Michigan, Institute for Social Research, 1975. (Appendix B to *Measures of political attitudes*, 1968)

Rogers, C. R. and Dymond, R. F. *Psychotherapy and personality change*. Chicago: The University of Chicago Press, 1954.

Rogers, C. R. The necessary and sufficient conditions for therapeutic personality change. *Journal of Consulting Psychology*, 1957, *21*, 95-103.

Rogers, C. R. A process conception of psychotherapy. *American Psychologist*, 1958, *13*, 142-149.

Rogers, C. R. A theory of therapy, personality and interpersonal relationships as developed in the client-centered framework. In S. Koch (Ed.), *Psychology: A study of science. Vol. 3*. New York: McGraw-Hill, 1959.

Schneider, K. [On the psychology and psychopathology of affective experiences.] *Zsch. f. d. ges. Neur. u. Psychiat.*, 1928, *2*, No. 1067.

Skinner, B. F. *The behavior of organisms; An experimental analysis*. New York: D. Appleton-Century, 1938.

Skinner, B. F. *Walden II*. New York: Macmillan, 1948.

Skinner, B. F. *Beyond freedom and dignity*. New York: Knopf, 1971.

Snygg, D. and Combs, A. W. *Individual behavior*. New York: Harper, 1949.

Warren H. C. *Dictionary of psychology*. Boston: Houghton Mifflin, 1934.

Wolman, B. B. *Dictionary of behavioral science*. New York: Van Nostrand Reinhold, 1973.

Wylie, R. C. *The self concept, volume 1: A review of methodological considerations and measuring instruments*. Lincoln: University of Nebraska Press, 1974.

NSPER: 76

Baltimore Presentations

Dr. David E. Price

Dr. David A. Payne

AFFECT: A BIOCHEMICAL PERSPECTIVE

David E. Price, M.D.
The Johns Hopkins University

I was asked to talk briefly about biomedical research that may be relevant to the subject of your meeting. The term "affective domain" was not a part of my vocabulary, so my first requirement was to understand your definition, and I was thus introduced to Bloom's *Taxonomy of Educational Objectives* (1956). There I found that the affective domain embraces, "objectives which emphasize a feeling tone, an emotion, or degree of acceptance or rejection. Affective objectives vary from simple attention to selected phenomena to complex but internally consistent qualities of character and conscience . . . a large number of such objectives in the literature expressed as interest, attitudes, appreciations, values and emotional sets or biases." One popular medical dictionary defines "affect" as, "a Freudian term for the feeling of pleasantness or unpleasantness evoked by a stimulus; also the emotional complex associated with a mental state; the feeling experienced in connection with an emotion."

The ideas expressed in the term "affect" are ancient and have inspired much of the world's classic literature. The anatomical location of the affect was long subject to some of the same uncertainties as the seat of the soul. Having wandered from bowels to heart to brain, the affect seems to have settled down and is now classified among functions ascribed largely to the nervous system.

Scientific knowledge about the affective functions has accumulated more slowly than has knowledge about the cognitive or motor functions of the nervous system. I suspect the reason is that science demands a rigor of identification, observation, measurement, and experimental manipulation that has come to this subject only in recent years and is still not highly developed. One obvious problem is the highly subjective nature of the manifestations of affect.

I would like to quote the closing paragraph of a lecture by Dr. Vernon Mountcastle (1975), a renowned neurophysiologist. Although not speaking specifically about affect, he has eloquently described the subjective nature of perception, which I consider to be a closely related neurologically mediated phenomenon. He said:

> Each of us lives within the universe -- the prison -- of his own brain. Projecting from it are millions of fragile sensory nerve fibers, in groups uniquely adapted to sample the energetic states of the world around us: heat, light, force, and chemical composition. That is all we ever know of it directly, all else is logical inference. Sensory stimuli reaching us are transduced at peripheral nerve endings, and neural replicas of them dispatched brainward, to the gray mantle of the cerebral cortex. We use them to form dynamic and continually updated neural maps of our place and orientation in the external world and of events in it. At the level of sensation, your images and my images are virtually the same and readily identified one to another by verbal descriptions, or common reactions. Beyond that, each image is conjoined with genetic and stored experiential information that makes each of us uniquely private. From that complex integral, each constructs at a higher level of perceptual experience . . . his own very personal view from within (p. 130).

The biomedical sciences have been concerned increasingly in the past decades with how the body works. With heavy emphasis on the anatomical sciences giving way to the study of physiology, we are delving ever deeper into the biochemical and biophysical processes of ever smaller constituents of the body and are now reaching down to the submolecular level.

Medical research is also marked by an increasing appreciation of the behavioral sciences and what they can contribute to the understanding of the human animal as a creature that functions socially as well as physically, intellectually and emotionally. One may question how independent these functions are. Perhaps they are only different manifestations of a complex chemically and physically mediated, biological system which has innumerable interrelated feedback pathways.

It is hard to isolate the affect for separate study because it appears to be influenced by a vast number of hereditary, psychological, nutritional, physiological, environmental, social, and situational factors. The feelings evoked by stimuli may well be determined by the sum of such influences. In my current state of ignorance, I find it difficult and perhaps futile to try to identify research that is particularly relevant to the affective domain. I am, therefore, going to mention several lines of investigation that may be of some interest and leave it to you to judge their relevance.

I asked one of my psychiatrist colleagues what he would say if you asked him to make this talk. Two lines of investigation immediately came to mind. One of these is the discovery of the catecholamine system of neurotransmitters for which Julius Axelrod received the Nobel prize in 1970 (Friedhoff, 1975). The discovery that noradrenalin and dopamine are interneuronal transfer agents in the brain and that they may be associated with affective disorders opened the way for study of the mechanism of action of neuropharmacologic agents and for an understanding of the role of drugs both in producing affective disturbances and in relieving them. Considerable attention has been focused on the use of such drugs as ritalin and dextroamphetamine as aids in dealing with hyperactive, underachieving children. As you know, there is controversy about the efficacy of this intervention and about the mechanism of any benefit that is observed. These drugs act upon the neurotransmitter system, and I mention them as justifying your interest in the expanding field of neuropharmacology.

The second avenue of progress he cited is the perfection of analog scales for measurement of affect. Although themselves not new, the modern application of analog scales is believed to have given behavioral scientists a method for exploring affective phenomena that is contributing much to improved understanding.

Discovery of the ultrastructure of chromosomes, the chemical nature of the gene, and the replication mechanism of the DNA molecule have all heightened interest in the field of genetics. Many abnormalities have already been identified that seem to be caused by the abnormality of a single gene that determines the presence of some important enzyme. I will mention two examples, one of which has important implications for intellectual development, the other for a digestive function. Perhaps both of these bear some relation to affect. One is the disorder popularly known as PKU or phenylpyruvic ketonuria caused by a deficiency of an enzyme which acts upon one of the amino acids. If not treated by appropriate dietary restriction in infancy and childhood, these individuals suffer severe intellectual retardation. Another common disorder is due to deficiency of the enzyme lactase which digests milk sugar. These individuals are intolerant to milk from which they suffer abdominal discomfort and diarrhea, a situation surely calculated to modify ones feelings about himself and the world around him.

Turning to the field of nutrition, one discovery from animal studies is that a protein deficiency in the mother's diet during gestation results in impaired behavioral development of the offspring which cannot be repaired after birth by supplemental feeding of the young. Studies are being conducted to determine whether the observation holds true among human mothers and their babies. If it does, supplementation of deficient diets of pregnant women may help assure normal behavioral development and the improvement of emotional well being.

There is not time to enumerate examples from among numerous environmental factors that clearly influence the affect: such things as physical comfort, noise, and toxic substances. However, I do not want to quit before mentioning infectious agents. One group of these has become known as slow viruses because of their propensity to lie dormant for many years after infection before producing a recognizable disorder. Some have their most characteristic action in the nervous system and the study of these viruses has become a fruitful field. It was less than three weeks ago that a Nobel prize was announced for Dr. D. Carleton Gajdusek who related one of these viruses to a neurological disorder transmitted among members of a primitive tribe by their practice of ritual cannibalism.

I think it is safe to say that the affect may respond to almost any disburbance of bodily well being, so that the whole gamut of medical research may have relevance. Our objective should be to achieve the highest possible state of health. I would therefore appeal for manipulation of the affective domain of educational objectives to attain that goal. Knowledge of healthful practices of daily living is not sufficient. Ways must be found to gain acceptance, internalization, and motivation, if that knowledge is to become effective as a determinant of behavior. We already know so much more than we apply. While many examples could be cited, consider only one. We know that for one common type of cancer, a potent preventive lies within the reach of all. Yet we have not discovered how to motivate people to forego cigarette smoking!

REFERENCES

Bloom, B. S. (ed.) *Taxonomy of educational objectives*. New York: David McKay Co. Inc., 1956.

Friedhoff, A. J. (ed.) *Catecholamines and behavior*, Vol. 2. New York: Plenum Press, 1975.

Mountcastle, V. B. The view from within: Pathways to the study of perception. *Johns Hopkins Medical Journal*, 1975, *136*, 109-131.

THE ASSESSMENT OF AFFECT: NOMOTHETIC AND IDIOGRAPHIC

David A. Payne
University of Georgia

The choice of the terms nomothetic and idiographic in the title was intentional, not capricious, and did not represent an attempt to cast an aura of greater wisdom and scholarship to these proceedings than already exists. Those terms represent at once the general thrust of my remarks *and* the cause of a partially unsolvable problem. The problem question being, "Does the simple aggregation of idiographic data result in a nomothetic summary?" The answer is probably both yes and no. "Yes" in a sense that in practice most applications of environmental measures are simply accomplished by summing the responses of a large number of respondents in affect assessment for an environment versus an individual person in that environment.

It is undoubtedly a perversion to use Allport's (1942) terms nomothetic and idiographic to simply describe a differential focus on general patterns of group behavior versus the individual. But this basic distinction will be used for expository purposes.

ON THE NATURE OF AFFECT AND THE AFFECTIVE MOVEMENT IN EDUCATION

There is a definite, strong, and pervasive revolution taking place in education. This revolution concerns what can best be described as affective learning outcomes. These outcomes are commonly concerned with such personal expressions as attitudes, interests and values. Evidence of this affective revolution can be seen in the type and extent of research on affective outcomes being completed and published in the professional journals. It is also evident in the papers being presented and discussed at professional educational and research meetings and conventions, in the kinds of sensitizing experiences being made part of teacher training programs, in the kinds of books on "humanizing the school curriculum" that are being published, and most of all in the actual learning experiences being implemented in our classrooms. Almost every teacher is aware that no matter what he or she does, affective learning takes place. Gagné (1965), for example, has noted that:

> . . . there are many aspects of the personal interaction between a teacher and his students that do not pertain, in a strict sense, to the acquisition of skills and knowledges that typically form the content of a curriculum. These varieties of interaction include those of motivating, persuading, and the establishment of attitudes and values. The development of such human dispositions as these is of tremendous importance to education as a system of modern society. In the most comprehensive sense of the word "learning," motivations and attitudes must surely be considered to be learned (p. 23).

Affective and cognitive characteristics are not separable. They develop together and influence together (Gordon, 1970). Concern with both kinds of outcomes, then, evidences concern for the total individual.

Douglas Heath, in an extremely informative article in the May 1972 issue of the *School Review*, has noted three major societal influences which are "fueling" the movement toward affective objectives in our schools. I draw heavily at this point upon his thoughts and ideas.

The first major influence reflects an awareness on the part of all citizens, but particularly those in charge of setting policy for our educational systems, of the apparent gulf between where society is going and our mode of life, and the way we are educating our youth. It appears to some observers that contemporary life can be characterized by the availability of an increasing (1) amount of leisure time, (2) degree of impersonalness in our interaction, (3) degree of interdependence, and (4) affluence. Perhaps todays youth are closer to perceived educational irrelevance than we were. Perhaps perceived and actual relevance are further apart than ever before in our history. Another contributor to the archaic character of traditional education is the changing role expectations society now holds for males and females.

A second major influence contributing to affective education of humanistic education movement stems from what might be called the emotional estrangement of youth. Heath suggests that, for example, the minimal demands of watching television, being basically a passive activity, contribute to this estrangement. Vicarious experience is one thing, direct self-initiated action and coping experience another. There would seem to be here an implication for a different kind of learning environment – an active, involved, hands-on set of experiences. Other fuels relate to sensory over-stimulation, changes in time perspective and perhaps lack of a sense of history, and ways of relating to authority. Perceived validity of one's self comes from the increased subjectivity that accompanies affective training.

The last primary impetus for affective education comes from the over-emphasis, perhaps thanks to some extent to Sputnik, on cognitive outcomes in education. We introduce more and complex subjects at earlier grade levels. Because of increased time spent on learning there is less time to feel. As noted previously concern for the total organism evidenced concern for both the head and heart. But what is this thing called affect?

THE MEANING OF AFFECT

To attempt to define the concept of affect or affectivity apart from a thousand or so other characteristics of human organisms is impossible at least and foolhardy at best. But what fools these psychometricians be to tread where angels fear. It is obviously necessary to get an operational handle on the construct of "affect" if we are going to attempt to measure it. Every speaker worth his salt at sometime or another employs the dictionary to help him initiate a topic or support a point of view. I would not want you to be disappointed. The always accurate and ever responsive lexicographers English and English (1958) in their *Comprehensive Dictionary of Psychological and Psychoanalytical Terms* note that the term *affect* at least historically, refers to a general class name for feeling, emotion, mood, and temperament. It is further emphasized that affect is generally *not* considered to exist alone apart from any cognitive dimensions of personality. Further, the term affect should not be construed as implying a unitary trait. Even a cursory perusal of the literature reveals the variety of ways that researchers have created to describe the affective domain. Table 1 contains a summary of three of these classifications. They represent kind of a mixed bag. But some commonalities exist. The terms "interests," "attitudes," and "values" are evident in two of the lists, and implied in the third (Krathwohl, et al. 1964). Even the authors of the *Taxonomy* imply in their writings that Receiving through Valuing include "interest," and that Responding through Organization includes "attitudes" and "values." Nunally (1967) uses really an old fashioned term "sentiments" as his collective noun rather than the term affect. He draws some interesting distinctions between the three terms Interests, Values, and Attitudes. He defines Interest as a preferance for an activity, Value as a broader life oriented goal, and Attitude as a feeling about particular objects (social, physical, or abstract). Are there any characteristics common to these three facets of affect? With the help of Shaw and Wright (1967), I began a list. Let me share it with you.

1) *Affective variables give rise to motivated behavior*. Feelings can motivate behavior; they can influence an individual to

TABLE 1
Alternative Ways of Conceptualizing the Affective Domain

Krathwohl, et. al. (1964)	*Center for the Study of Evaluation (Hoepfner, et. al. 1972)*	*Nunally (1967)*
Receiving Responding Valuing Organization Characterization by a Value or Value Complex	Personal Temperament Social Temperament Attitudes, Opinions, Beliefs Needs Interests Values	Sentiments Interests Values Attitudes

respond in different ways, and different individuals to respond differentially. The concept of individual differences is necessary, but not sufficient, evidence for the existence of an affective variable.

2) *Affective variables vary in intensity*. Two individuals may have favorable attitudes toward the same referent, but vary in the intensity of their feelings. Conversely, they may hold different attitudes with the same degree of intensity. It is probably true that motivational strength is tied to intensity of feeling and, therefore, that the likelihood of certain behavior varies with intensity. Intensity also has implications for instruction. The more intense an attitude, the harder it is to change it.

3) *Affective variables are learned*. The cumulative effect of training, education, childrearing, and formal and informal social interactions influence the development of attitudes, interests, and values.

4) *Affective variables have specific referents*. These referents need not always be concrete objects, but may include abstract or social referents related, for example, to world or political issues or to theology.

5) *Affective variables represent varying degrees of interrelatedness*. Attitudes toward similar objects are more likely to be interrelated than attitudes toward dissimilar objects. Complex clusters of interrelated beliefs (e.g., toward the women's liberation movement, and its subissues of abortion, and equal pay for equal work) are more difficult to change than single, narrowly referenced attitudes. This dimension is sometimes referred to as "generality."

6) *Affective variables are relatively stable and enduring*. The history of reinforcement of an attitude in a particular individual is probably the primary determinant of the stability of the attitude. Affective predispositions are difficult to change. The role of family and school – the primary social institutions – in developing and modifying attitudes cannot be underestimated.

7) *Affective variables vary in salience*. Salience refers to proximity to the surface of a person's mind – how easily an emotion can be evoked.

I sat back at my desk somewhat smugly and looked at the list. It then dawned on me that these descriptors could be used with any human characteristic, cognitive, affective, and to a great extent psychomotor. What then Horatio, is there nothing new under the psychometric sun?

THE NEED TO ASSESS AFFECTIVE CHARACTERISTICS

There are four primary reasons why affective outcomes need to be dealt with and assessed in our educational institutions:

1) *Affective variables influence an individual's ability to participate effectively in a democratic society.* Attitudes toward institutions, practices, social groups, and the like, affect and are affected by the efforts of society to maintain itself and meet the needs of its members. If for no other reason than this, affective objectives must be considered legitimate outcomes of concern to educators (Scriven, 1966).

2) *The development of skills and abilities related to the acquisition and growth of attitudes and values is necessary for a healthy and effective life.* The development of rational attitudes and values is the result of intelligent examination of society's needs and those of the individual. Affective skills are necessary to the overall effective functioning of the individual in society.

3) *Affective outcomes interact with occupational and vocational satisfaction.* In maintaining himself economically, an individual must (a) relate effectively with his associates, (b) enjoy his work, (c) believe it possible to make maximal use of his abilities, and (d) feel that he is making a contribution to society. Kahl (1965) reasons that the values of mastery, activism, trust of others, and independence of family should be considered legimate educational objectives, since they have been empirically related to socioeconomic achievement and upward mobility in our heavily industrialized society.

4) *Affective variables influence learning.* This postulate has been well documented. The interaction of teachers' and students' affective characteristics influences progress toward the attainment of classroom goals. Ripple (1965), in summarizing research on the affective characteristics of the learning situation, concluded that the attainment of classroom objectives is facilitated by (a) a

generalized feeling of warmth in the learning environment, (b) tolerance of emotional and feeling expressions on the part of students, (c) democratic group decision-making leading to stimulating activities, (d) the use of nonpunitive control techniques of considerable clarity and firmness, (e) reduced frustration and anxiety in learning tasks, and (f) shifting states of order based on the organization of emotions toward the achievement of goals. More specifically, Domino (1971) has experimentally demonstrated interaction between a student's achievement values, the instructor's teaching style, and the amount of, and satisfaction with learning. If students are learning materials that interests them, they are likely to develop positive attitudes toward it. Attitudes have also been shown to be related to achievement. Bassham, Murphy, and Murphy (1964) have demonstrated with a sample of sixth-grade students a relationship between positive attitudes and achievement in arithmetic (see also Aiken, 1970).

A cautionary note needs to be added. The desire to improve, modify, adjust, expand, or in some way influence and alter attitudes and values should not obscure our primary concern which is with learning. We do not *just* want to make the students feel better. Heath (1972) suggests that " . . . we need to educate youth, not just his head nor his heart. The promise of affective education is that it will stimulate us to recover the person lost among our abstractions; its danger is that it may devalue man's most promising adaptive and educable skill: a disciplined intellect" (p. 371).

But our concern here initially is with educational environments rather than individuals. What is the connection between individual affect and the dimensions of educational environments. Let us step back in time and take a brief historical sighting by considering theorists who have attempted to reconcile the individual and the group.

IN THE BEGINNING

Many believe that a good researcher must always begin with a theory. Educational researchers have been using many theories in their search for variables which relate to student achievement or performance in the school. The seemingly unending search is for those variables which can account for the most variance so that eventual control and manipulation can be exercised to enhance and maximize student growth.

In the beginning there was Lewin (1936) and his field theory which postulated that behavior was the result of the interaction of two independent vectors: person and environment. The famous postulate that behavior was a function of person and environment interacting, B = F (P,E,) sent many a researcher scampering with his calipers and magnifying glass. Henry Murray's (1938) "personology" with the need-press dimensions also contributed significantly to the "environmental" literature. More contemporary theorists are now having an impact on the kinds of research being undertaken. Illustrative of the newer theorists would be Barker (1968) and his ecological psychology, Holland (1966) whose theory is embeded in a vocational decision-making and satisfaction framework, and Pervin's (1967) work with his technique called Transactional Analysis of Personality and Environment which is basically a six concept, 11-point, 52 scale semantic differential. A recent and important critique of five major theoretical positions with regard to person-environment interaction has been presented in a major monograph by Walsh (1973).

Recent research on the school environment has focused on testing hypotheses derived from a sociopsychological theory of the classroom as a social system (Getzels and Thelen, 1960). This particular model suggests that institutional and individual characteristics interact in classrooms to influence school learning. Getzels and Thelen incorporated the views of both Lewin and Murray by stating that the social system and resultant behavior were a function of the "simultaneous" interaction of the nomothetic and idiographic dimensions:

> . . . social behavior results as the individual attempts to cope with an environment composed of patterns of expectations for his behavior in ways consistent with his own independent pattern of needs (p. 132).

What Getzels and Thelen have essentially done is to take the environment and person dimensions used to study personality by Lewin and Murray, and expand them to the concept of an entire social system such as the school (Fig. 1). Such an extension is not

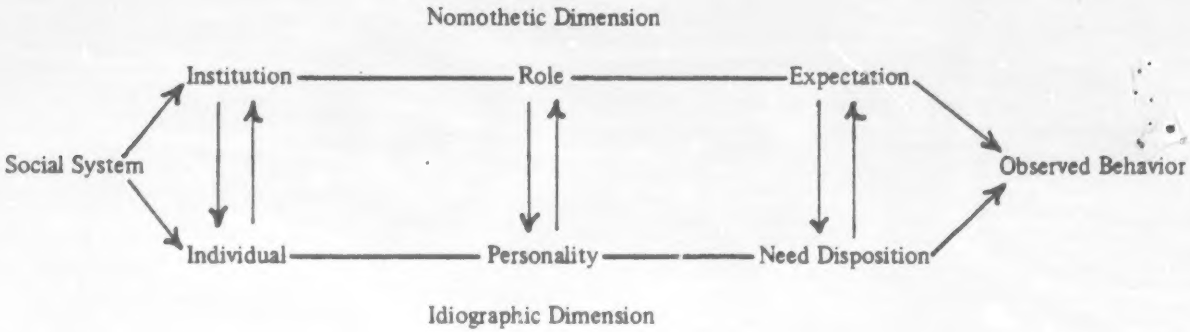

Nomothetic Dimension
Institution
Role
Expectation
Social System
Observed Behavior
Individual
Personality
Need Disposition
Idiographic Dimension

Figure 1: Getzels and Thelen's model of an Interactive Social System

unreasonable. As Shoben (1962) points out, the school is more than just a place to learn and develop academic skills. It is essentially a mini-community where members interact and influence themselves and others.

Coughlan and Cooke (1964) have, in addition, suggested that teacher attitudes and perceptions of the school environment are related to student performance and the effectiveness of the school. Several other investigators have looked at the influence of classroom social climate and students' and teachers' characteristics on learning. The findings in general suggest that affective aspects of classroom climate (satisfaction, intimacy and friction) predict both cognitive and affective learning. With respect to affective learning outcomes, Walberg (1969) has reported significant multiple correlations of .44 and .41 between 14 dimensions of the classroom social climate, and science interest and physics interest, respectively. Thus it appears the environmental characteristics constitute viable variables potentially useful in educational research focused on student achievement or performance in the school.

EVIDENCE OF THE IMPORTANCE OF ENVIRONMENTAL VARIABLES

The literature abounds with reports bearing on the significant impact of environmental and climate factors on the teaching/learning situation. Although many studies show significant correlational relationships between environment and outcome measures, very few experimental studies can be found where independent variables were systematically manipulated under controlled conditions. In any event, the views of Anderson and Walberg (1974) and Randhowa and Fu (1973) strongly support the environmental line as one to follow in educational research. Anderson and Walberg (1974), for example, cite data which support the superiority of environmental measures over I.Q. scores as predictors of cognitive, affective, and behavioral criterion measures.

After a very extensive review of the relevant literature, Perkins (1976) was able to draw several conclusions. Among the more prominent were the following:

1) High inference measures of the school environment are generally more meaningful than low inference measures in educational research, and high inference measures are especially appropriate in studies concerned with the sociopsychological learning environment. To suggest the lack of validity for low inference measures is almost heretical. After so much work has been done, time and effort expended in creating and applying observational techniques, for example, we really are not closer to the identification of critical variables or the development of valid measures than we were 40-50 years ago. The use of observation systems be they category or sign has resulted in the fragmentizing of behavior to the point where it loses meaning. In addition, the inferences that must be made from an isolated behavioral act to a motivational referent, particularly in regard to affective variables, are extensive and intensive, and probably not warranted by the data.

2) Students perceive differences in the learning environment related to the course content being studied. Science classes were perceived as more rigorous, formal, and fast-paced groups with moderate goal direction and difficulty, while humanities courses were seen as being slower-paced and more disorganized.

3) Students' perceptions of the classroom learning environment do affect student performance and achievement. Students' perceptions of the difficulty of the class were shown to relate positively to achievement while perceived environmental dimensions such as friction, apathy and cliqueness were found to be inversely related to achievement.

4) While the variables of class size and teacher personality do affect students' perceptions of the classroom learning environment, the variable of sex of teacher does not. As classes increased in size, students

> perceived the environment as more formal and diverse and less difficult and intimate. Personality characteristics of teachers, such as self-centeredness and authoritarianism, were found to be related to the perceived organization, formality, supervision, and animosity in the classroom environment. It also appears that attitudes toward school appears to be negatively related to grade level and positively related to socio-economic class.
>
> 5) While little research directly related to teachers' perceptions of the school learning environment exists, early related research focused on simply identifying factors affecting teacher job satisfaction and developing measures of teacher work values. Only recently has an attempt been made to assess the relationship between teachers' perceptions of the school environment and student performance. Teachers' perceptions of the school environment with regard to dimensions such as student evaluation practices, community relations, opportunities for faculty development, and faculty contribution to the educational program were shown to be positively related to school performance (pp. 34-35).

METHODOLOGIES INVOLVED IN ENVIRONMENTAL ASSESSMENT

Menne (1967) has identified two major variables that are involved in the assessment of any learning environment. These are source of data and nature of data. By specifying two major data sources (the individual student as opposed to institutional data) and two types of data (psychological and perceptual as opposed to objective) a general model can be developed which describes the major approaches to environmental assessment. The model might look something like the one shown in Fig. 2.

The first type of measure Menne (1967) has characterized as *subjective student perceptions* (Quadrant A). This class includes a myriad of self-report perception instruments based on the assumed validity of "high inference" measures, i.e., when general

		DATA SOURCE	
		Student	Institution
NATURE OF DATA	Perceptions & Feelings	A	C
	Objective Behavioral Reports	B	D

Figure 2: Model for describing major approaches to environmental assessment

or global impressions and/or opinions are reported. Such methodologies usually require students to rate a variety of dimensions of the environment. Typical would be the Learning Environment Inventory (Anderson, 1973). The LEI contains 14 scales (e.g., Cliqueness, Difficulty) appropriate for secondary level students. Each respondent indicates, using a four point scale, his or her agreement or disagreement on how well each of 105 item/statements describe a particular class (e.g., "Many students in my school would have difficulty doing the advanced work of my class"). This method of measuring an educational environment was developed for college level applications by Pace and Stern (1958) and refined by the late Dr. George Stern as reported in his very significant volume entitled *People in Context* (1970). A brief vignette from the 1958 report by Pace and Stern should provide the reader with the flavor of the kinds of profiles that can be drawn using their instrument, the College Characteristics Index:

> The total picture of the environment, then is one of high social activity, esprit de corps, and enthusiasm combined with an emphasis on helping others and idealistic social action and all within a fairly well understood set of rules and expectations which are deliberative and orderly. One would expect some of the explicit objectives of such an institution to stress personal and social development, idealism and social action, and civic responsibility (p. 274).

This description of the "institutional press" can be contrasted to the institutional intent and prospective college student's personal needs. Such information should prove useful to college administrators, high school counselors, and any who are involved in college admissions work in matching individuals and institutions.

The approach suggested in Quadrant B to the measurement of environmental characteristics relies on reports of specific observable self-reported student behaviors. The *observable behaviors* might range from number of hours spent studying per week to number of social activities. Astin (1965) has reported a study using a 35-item student self-report instrument with such statements as: "The instructor called students by their first names," and "I took notes regularly in class." The investigator was able to demonstrate significantly different classroom environments in different fields of undergraduate study. Another example of how the behavioral approach has become operationalized is the study of the environment of the classroom of gifted students, as reported in the work of Steele, House, and Kerins (1971), with their Class Activities Questionnaire.

Quadrant C suggests the use of other individuals (other than the student himself) to report on the impressions of the institutional environments. A great variety of observational systems are available for application (Simon and Boyer, 1970). The problem with this approach is that observational systems focusing on small segments of behavior compound the inference problem.

The last of our methodologies we will label, for lack of a better rubric, *institutional description* (Quadrant D). This method, pioneered by Astin and Holland (1961), focuses on objective and readily available data such as number of students, student/faculty ratio, number of volumes in library, tuition, percentage of males and females, distribution of degrees held by faculty, operating budget per student, and similar indicators. The types of data just described are probably more meaningful at the institution level rather than at a lower level, e.g., the classroom. The descriptions are also probably more meaningful at the college level, although some commonality with other educational levels obviously exists. Typical of the *institutional description* instruments is the Environmental Assessment Technique (EAT) (Astin and Holland, 1961) which is based on eight attributes of the student body: size, intelligence level of the students, and the "personal orientations of the students as indicated by the percentage of students in each of six classes of major fields (Realistic, Intellectual, Social, Conventional, Enterprising, and Artistic). Most of these data can be secured from readily available published sources. Illustrative of the kinds of data presented by Astin in support of the validity of the EAT is a 1963 study showing strong relationships between his EAT attribute scores and student descriptions of the perceived effects of four years of college.

ILLUSTRATIVE ENVIRONMENT CLIMATE MEASURES

This section of our presentation will deal with the brief descriptions of four instruments.

Class Activities Questionnaire

The first of these is the Class Activities Questionnaire (CAQ). Following is a brief description of the CAQ taken from the writing of Steele, House and Kerins (1971):

> The Class Activities Questionnaire (CAQ) is a 25 item instrument administered to both students and teachers. It asks students to agree or disagree on a four point scale to statements describing general kinds of activities which characterize their class. These activities imply either levels of thinking of affective classroom conditions. Each item is paired with another item to compose a factor; sixteen factors yield a revealing profile of the class. (Five factors are represented by single items. One factor, "Teacher Talk" is reported separately as well as being used as a component of the "Lecture" factor.) In addition, subscores are derived by clustering factors into the four dimensions of Lower Thought Processes, Higher Thought Processes, Classroom Focus, and Classroom Climate. The cognitive dimensions of Lower and Higher Thought Processes represent a dichotomy strongly supported in validation studies of Bloom's *Taxonomy of Educational Objectives* (Bloom, et al., 1956). The Classroom Focus dimension assesses whether the focus is on the teacher as information-giver with students having a passive role in the class. The Classroom Climate dimensions assesses attitudes and feelings, such as how relaxed and open the class is and the amount of involvement of students in class activities (p. 450).

A further description of the major elements of the CAQ is presented in Table 2.

Dimensions	Factors	Brief Descriptions
I. Lower Thought Processes	1. Memory:	Activities calling for recall or recognition of information presented.
	2. Translation:	Activities calling for paraphrasing or expressing information in a different symbolic form.
	3. Interpretation:	Activities calling for recognition of relationships and seeing implications of information.
II. Higher Thought Processes	4. Application:	Activities calling for selection of appropriate methods and performance of operations required by problem situations.
	5. Analysis:	Activities calling for recognition of the structure of material, including the conditions that affect the way it fits together.
	6. Synthesis:	Activities calling for the generation of new ideas and solutions.
	7. Evaluation:	Activities calling for development and application of a set of standards for judging worth.
III. Classroom Focus	8. Discussion:	Student opportunity for an involvement in class discussion.
	9. Test/Grade Stress:	High pressure to produce teacher selected answers for a grade.
	10. Lecture:	Teacher role is information-giver with a passive, listening role for students.
IV. Classroom Climate	11. Enthusiasm:	Student excitement and involvement in class activities.
	12. Independence:	Tolerance for and encouragement of student initiative.
	13. Divergence:	Tolerance for and encouragement of many solutions to problems.
	14. Humor:	Allowance for joking and laughter in the classroom.
	15. Ideas:	The extent to which ideas studied are valued over grades.
	16. No Enjoyment:	The extent to which students do not enjoy the ideas studied in class.

TABLE 2: Class Activities Questionnaire

The Learning Environment Inventory

The Learning Environment Inventory (LEI) is an instrument designed to measure the social climate of learning of a class as perceived by students. Its best application would be at the secondary school level. Developmentally, it is an expansion and improvement of the older Classroom Climate Questionnaire and describes the nature of interpersonal relationships in class, as well as the structural characteristics of a class as perceived by students. The LEI has been developed to include items representing scales indicative of concepts identified through research as good predictors of learning and consistent with sound social-psychological theory.

The LEI consists of 105 statements descriptive of typical school classes and requires the student to express his agreement or disagreement with each statement on a four point scale. The 15 scales of the LEI with representative sample items are listed below:

Scales	*Sample Items*
1. Cohesiveness	"Members of the class are personal friends."
2. Diversity	"The class divides its efforts among several purposes."
3. Formality	"Students are asked to follow a complicated set of rules."
4. Speed	"The class has difficulty keeping up with its assigned work."
5. Environment	"The books and equipment students need or want are easily available to them in the classroom."
6. Friction	"Certain students are considered uncooperative."
7. Goal Direction	"The objectives of the class are specific."
8. Favoritism	"Only the good students are given special projects."
9. Difficulty	"Students are constantly challenged."

10. Apathy . . . . "Members of the class don't care what the class does."
11. Democratic . . . . "Class decisions tend to be made by all the students."
12. Cliquenes . . . . "Certain students work only with their close friends."
13. Satisfaction . . . . "Students are well-satisfied with the work of the class."
14. Disorganization . . . . . "The class is disorganized."
15. Competitiveness . . . . . "Students compete to see who can do the best work."

In terms of its structure, the LEI can be used to derive scores for individuals within sampling units or a mean can be used to provide an estimate of the climate profile of a class, grade, or school.

Data generated from research with the LEI in the evaluation of the Harvard Project Physics in 1969, using some 64 classes and 1,048 students (Anderson, 1973), shows that test-retest reliability coefficients for the test scales vary from a low of .43 (Diversity) to a high of .73 (Friction). Coefficients of internal consistency, indicative of the degree to which items on the same scale tend to measure the same thing, vary from .54 (Diversity) to .85 (Goal Direction). The intraclass correlation, a class coefficient indicating group reliability of the LEI scales, varys from .31 (Diversity) to .92 (Disorganization).

Validity data supportive of the capability of the LEI scales to predict a variety of school outcomes stems from many studies (Anderson, 1970; Walberg and Anderson, 1968; Anderson and Walberg, 1968). In general, research shows that the subscales are predictive of learning outcomes, or other variables described as being theoretically conducive to the support of adequate learning environments. These relationships are, however, complex. The Cohesiveness scale, for example, relates to three major class and course properties. Small classes are more cohesive than are large classes, particularly when the class contains fewer than 16 pupils (Walberg and Ahlgreen, 1970). Class cohesiveness relates to learning criteria differentially, depending upon the norms of the cohesive class. Cohesive classes seemingly sanction only goal-directed behavior. If the group norm includes learning, cohesiveness contributes to increased learning. In non-learning oriented classes, cohesiveness acts against those pupils who want to learn (Anderson, 1970).

The My School Inventory

As originally developed, the My Class inventory contained 45 items distributed over 5 scales labeled Satisfaction, Friction, Competitiveness, Difficulty, and Cohesiveness. The inventory is considered appropriate for students aged 8 through 12. Reading level of My Class inventory items is generally considered appropriate for fourth grade students.

An adaptation of the My Class inventory was achieved by rewording My Class scale items to incorporate the concept of "school" rather than "class" or "classes." The modified instrument was called the My School inventory. For example, the word "class" was changed to "school" in the following inventory item: "The pupils enjoy their schoolwork in my class." This change was considered necessary in order to elicit a more generalized response about school characteristics, and to better take into account situations where individual elementary students had different classes with different teachers during the regular school day.

Individual scale reliabilities reported in the manual for the My Class inventory range from .54 to .77, considerably lower than those for the LEI. The instrument is still undergoing development, though it has been used successfully in several research and evaluation studies (Walberg, 1969; Cayne, 1970; Walberg, Sorenson and Fishback, 1972; Perkins, 1976). Perkins (1976) in a canonical correlational study using the My School inventory reported median internal consistency coefficients for 42 elementary schools range from .41 (Competitiveness) to .73 (Satisfaction). In an evaluation of the My Class inventory, Maguire, Gaetz, and Manos (1972) state: ". . . the My Class Inventory appeared to be superior . . . My Class produced more interpretable results . . . it provides a broader coverage of the entire atmosphere, in particular a greater effort is made to tap the interstudent relationships. For these two reasons, the My Class inventory is commended for use in the evaluation of primary grade atmosphere.

Sample items and the labels for the five My School scales are presented below:

Scales	*Sample Items*
1. Satisfaction.	"The pupils enjoy their school work in my class."
2. Friction	"Children are always fighting with each other."
3. Competitiveness.	"The same people always do the best work in our class."
4. Difficulty	"In our class the work is hard to do."
5. Cohesiveness	"My best friends are in my class."

Each scale contains nine items requiring a Yes or No response. Due to scoring program, high scores are interpreted in directions opposite that of label, e.g., Satisfaction = Dissatisfaction.

The School Survey

The teacher is another source of valuable descriptive data related to the environment. To assess teacher perceptions of various factors mediating the school environment, the School Survey may be employed. This instrument was developed by Coughlan and Cooke (1974) at the University of Chicago to assess a global construct called work attitudes or satisfaction of teachers within the school setting. They describe the instrument as being useful for the formative and summative evaluation of schools and for providing school personnel with feedback concerning teachers' work attitudes.

The School Survey contains 118 items each requiring one of three possible responses, Agree, ?, and Disagree.

Factor analysis of data from the third form of the School Survey (Coughlan, 1970) initially yielded 13 factors, each containing from 6 to 10 items and accounting for relatively moderate amounts of the total test variance. Kuder-Richardson Formula 20 internal-consistency reliability coefficients for individual factor scores range from .44 to .80 with a median of .67. Perkins (1976) has reported median internal consistency reliability coefficients for 42 schools ranging from .58 (School Community Relations) to .85 (Supervisory Relations). A more detailed discussion of research and development efforts with the School Survey can be found in Coughlan and Cooke (1974, pp. 295-313). Following are the scale labels for the now finalized 14 separate dimensions of the School Survey, together with a sample item from each scale:

Scales	*Sample Item*
1. Administrative Practices	"The administration seems willing to give careful consideration to our ideas and suggestions."
2. Professional Work Load	"I am asked to spend too much time in meetings around here."
3. Non-Professional Work Load	"I receive sufficient clerical assistance to do my job effectively."
4. Materials and Equipment	"The instructional materials provided for me here are very satisfactory."
5. Buildings and Facilities	"The layout of this school is inconvenient for the staff."
6. Educational Effectiveness	"In my opinion, adequate educational standards are being upheld in this school."
7. Evaluation of Students	"Students' absences are excessive in this school."
8. Special Services	"Our library services for students are very satisfactory."
9. School-Community Relations	"In general, I approve of school board policies."
10. Supervisory Relations	"I seldom get the help I need in handling difficult discipline cases."
11. Colleague Relations	"People in my school cooperate well."

12. Voice in Educational Program . . "I should have a greater voice in selecting students' textbooks and reference material."
13. Performance and Development . . "I think my work performance is judged fairly here."
14. Financial Incentives . . . . "I feel our salary system adequately regards outstanding work."

ILLUSTRATIVE SAMPLE DATA FROM THE LEARNING ENVIRONMENT INVENTORY, MY SCHOOL INVENTORY, AND SCHOOL SURVEY

As part of a larger Title III project concerned with the assessment of principal competencies (Payne, et al., 1975), the LEI, My School, and School Survey instruments were administered to ten school samples of secondary students (N=3613), 35 elementary school student samples (N=3350), and a sample of 35 teacher/school unit samples (N=1145). School means were used as the unit of analysis. The dependent measures for the elementary students were: Vocabulary, Reading, Language, Work Study, and Math grade equivalents from the Iowa Tests of Basic Skills. For the secondary students, the Reading and Math scores from the Tests of Academic Progress were used. Average daily attendance figures for the 20-day school period closest to the testing dates were also used as criterion measures for both samples.

Relationships Between School Survey Factors and Student Achievement and Average Daily Attendance

Table 3 presents the intercorrelations between School Survey factors, student achievement, and average daily attendance for the elementary school sample. For this sample, 86% (72 out of 84) possible achievement correlations were statistically significant. The rather large number of significant correlations for the elementary sample points out the strong interrelationship be-

TABLE 3
Intercorrelation Between School Survey Factors (Teacher Satisfactions) and Student Achievement and Average Daily Attendance
(Elementary Schools: N=35)

School Survey Factors	Iowa Tests of Basic Skills Subtests						% Average Daily Attendance[2]
	Vocabulary	Reading	Language	Work Study	Math	ITBS Total	
1. Administrative Practice (9)[1]	.49*	.50*	.49*	.55*	.49*	.53*	.27
2. Professional Work Load (9)	.40*	.44*	.42*	.45*	.43*	.43*	.27
3. Non Professional Work Load (6)	.27	.31	.28	.31	.28	.31	.27
4. Materials and Equipment (8)	.21	.23	.20	.20	.25	.22	.22
5. Buildings & Facilities (7)	.47*	.50*	.42*	.42*	.38*	.49*	.25
6. Educational Effectiveness (10)	.81*	.82*	.83*	.82*	.76*	.80*	.55*
7. Evaluation of Students (10)	.68*	.70*	.69*	.73*	.67*	.67*	.57*
8. Special Services (8)	.48*	.48*	.43*	.49*	.47*	.45*	.13
9. School Community Interface (7)	.49*	.47*	.50*	.53*	.57*	.51*	.08
10. Supervisory Relations (10)	.37*	.41*	.41*	.36*	.39*	.33*	.19
11. Colleague Relations (7)	.57*	.55*	.54*	.56*	.59*	.54*	.31
12. Voice in Educational Programs (8)	.46*	.43*	.41*	.44*	.38*	.41*	.27
13. Performance and Development (10)	.35*	.35*	.36*	.40*	.40*	.35*	.14
14. Financial Incentives (9)	.41*	.40*	.37*	.47*	.41*	.45*	.09

[1] Number of items on scale

[2] Based on a sample of 36 schools

*Correlation significantly different from zero, $p < .05$

tween various subtests of the Iowa Tests of Basic Skills. Supplemental analyses of the ITBS revealed high positive correlations approaching .95 between these achievement indices. The data for elementary schools do, however, point out the predictive relationship between the School Survey factors as measures of teachers' job-related attitudes and student achievement.

Of particular interest in Table 3, for the elementary sample, are the School Survey scales of Educational Effectiveness (6) and Evaluation of Students (7). The intercorrelations between these scales and the ITBS subtests show a rather strong association between teachers' perception of the school's general educational effectiveness, as regards academics and curricula, and its programs, procedures, and materials for adequately evaluating students and elementary student achievement scores. The correlations would predict that teachers having poorer attitudes towards these dimensions of their working environment are found in schools with relatively lower student achievement. These same two School Survey scales (6 and 7) were the only scales to correlate significantly with average daily attendance. What might be represented in the data is the global nature of the schools in the sample and their teacher/student/school environment characteristics. It is reasonable to assume, for example, that better school systems, having more effective school programs, are generally associated with better teacher work attitudes, higher average daily attendance, and thus, higher student achievement.

It appears that the relationships between the School Survey, as a measure of teacher attitudes "mediating" the school environment, and elementary student achievement are rather strong in terms of the frequency of significant correlations, though the magnitude of relationships for the various scales varies from moderate to high.

Relationships Between the My School and Learning Environment Inventory and Student Achievement and Average Daily Attendance

The data derived from correlating the school climate/learning environment measures with outcome variables were somewhat supportive of predicted strong relationships between these two sets of variables. The relevant intercorrelations for elementary and secondary schools are presented in Tables 4 and 5.

For the elementary sample, 31% of the 35 correlations were found to be statistically significant. For the secondary schools (Table 5), 22% of the 45 possible correlations were statistically significant. The magnitude of the significant correlations for the secondary sample was somewhat higher than that for the elementary sample. Owing to the rather high intercorrelation between the ITBS subtests mentioned previously, the most meaningful correlations for elementary schools are those between the ITBS Total score and average daily attendance, and the My School Difficulty (read "Easiness") and Cohesiveness (read "Lack of Cohesiveness") scales. Apparently the degree to which elementary students see the school and school climate as *not* being overly demanding and loosely organized and structured in terms of interpersonal relationships with others is positively related to their achievement test performances. The scores on the Difficulty factor of the My School are also related to average daily attendance in this sample. In a comparison of the magnitude of the correlations between the Difficulty and Cohesiveness scales and the ITBS Total scores, the Difficulty factor accounts for approximately three and one-half times as much variance in student achievement as the Cohesiveness scale, and thus, is considered the better scale for prediction purposes.

In comparing data from Tables 4 and 5, some similarities between elementary and secondary students' perceptions of school climate/learning environment and student achievement exist. Secondary students' perception of a somewhat loosely organized (Scale 14, Disorganization), less goal directed (Scale 7, Goal Direction), but challenging (Scale 9, Difficulty) atmosphere tends to be rather strongly associated with student achievement as measured by the Tests of Academic Progress. Of interest, when comparing the two samples, is the magnitude and direction of the Difficulty/student achievement correlations. Apparently secondary students' perception of the learning environment as *more* difficult is associated with higher student achievement. For elementary students, the trend is reversed.

Another interesting finding for the secondary schools was the significant positive correlation between LEI scale 10 (Apathy) and Reading, and the negative correlation between Apathy and

TABLE 4

Intercorrelation Between My School Factors and Student Achievement and Average Daily Attendance (Elementary Schools, N-35)

My School Factors[1]	Vocabulary	Reading	Language	Work Study	Math	ITBS Total	% Average Daily Attendance[2]
	Iowa Tests of Basic Skills Subtests						
Satisfaction	.21	.22	.13	.23	.14	.19	.08
Friction	.18	.18	.26	.21	.31	.19	.30
Competitiveness	.11	.13	.14	.11	.21	.13	.04
Difficulty	.65*	.60*	.65*	.66*	.63*	.65*	.55*
Cohesiveness	.35*	.34*	.31	.37*	.28	.35*	.27

[1] Due to scoring procedure applied, scale interpretation should be in direction opposite that listed (e.g., Satisfaction = Dissatisfaction)

[2] A.D.A. correlation based on N = 36 Elementary schools

*Correlation significantly different from zero, $p. < .05$

TABLE 5

Intercorrelation Between Secondary Student Achievement and Average Daily Attendance and Learning Environment Inventory Factors (Secondary Schools, N=10)

Tests of Academic Progress	Learning Environment Inventory Factors														
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*
Reading	−.31	.43	−.49	.58	−.51	.35	−.71*	.62	.93*	.65*	.58	.46	.61	−.86*	−.55
Math	−.34	.41	−.53	.60	−.53	.29	−.74*	.60	.89*	−.65*	.58	.48	.62	−.85*	−.61
School Average Daily Attendance	.75*	.72*	.44	−.40	.59	.03	.46	.20	.12	.54	−.04	.16	−.48	.32	.41

*Correlation significantly different from zero, $p < .05$

Math. Such a finding may suggest something about the differential attitudes of high and low achieving students in Math and Reading and their relative perceptions of the amount of apathy in the general learning environment. High perceived apathy is *positively* associated with high reading achievement, and *negatively* associated with achievement in math for these schools. An additional variable, possibly useful in explaining these data, is the different instructional approaches teachers take to these two subjects.

An additional finding of interest for the secondary schools are the correlations between the LEI scales of Intimacy (1), and Diversity (2), and average daily attendance. Apparently, for this sample, students' perception of a school environment as offering warm relationships with others, and providing a multitude of activities and academic opportunities that match student interests is rather highly and positively related to average daily attendance. When considering adolescent enjoyment of the school environment as a factor in fostering good school attendance, these scales and their correlations seem quite logical.

In summary, it appears that students' perception of a general school learning environment, characterized by an easygoing, unpressed, somewhat structured, diverse and warm atmosphere, makes the greatest contribution to student learning as measured by standardized achievement tests. In addition, student achievement for both elementary and secondary schools seems predictable from students' assessment of the relative easiness (elementary) or difficulty (secondary) of the school environment.

The Use of Environmental Data to Build Pre-Service and In-Service Training Programs

In addition to the research applications that were just noted, an additional application of environmental data can be illustrated. Data from the previously described School Survey will be used.

The illustration is drawn from the principal competency project previously noted (Payne, et al., 1975). As part of that project, individual school principals received a profile of their school with respect to the 14 scales of the School Survey. A feedback form similar to that presented in Table 6 was used.

Individual school administrators could then identify the "highs" and "lows" of the working environment as perceived by the teachers. The aggregate data of Table 6 present a relatively flat profile, but it appears that highly rated were elements in the school environment labeled Professional Work Load, Colleague Relations, Non-Professional Work Load, and Supervisory Relations. Less highly rated were Administrative Practices, Voice in Educational Programs, and not unexpected, Financial Incentives. Principals used their individual school profiles to develop a set of objectives that became part of a field-based training program being conducted by Valdosta State (GA) College under the supervision of Dr. Joseph Licata.

The same general approach could be used to develop programs for teachers.

PROBLEMS IN DEVELOPING AND APPLYING CLIMATE/ENVIRONMENT MEASURES

Menne (1967) and others have commented on the many problems, both theoretical and technical, that must be faced at the outset of the development of environment measures. Among the dominant problems are those related to:

1. *Content Validity*: As is the case with the development of any instrument, the domain of phenomena to be measured must be clearly specified. The developer must ask himself *why* a particular variable or set of variables are to be measured. Is it for purposes of institutional comparisons, program evaluations, or to assist an administrator or teacher in defining the parameters of the dominant climate so that modifications could perhaps be made? It is important to emphasize that the focus of the instrument is on the environment not the inhabitants. It follows that the definition of broad variables, e.g., liking school, does not do much to clarify a complex phenomena.

2. *Sampling*: At both the pilot or field-test stages where data are gathered for instrument refinement purposes and at the stage of gathering final data, great care needs to be given sample selection. Matrix sampling procedures can be used in a way similar to that described by Walberg and Welch (1967). The over-all experimental design used to validate the instrument might employ constrasted groups. Again in such a situation identifica-

TABLE 6

Summary of Means and Standard Deviations for the Dimensions of the School of the School Survey Administered to 1200 Teachers in Forty-Five Schools*

	Number of Items on Scale	*Assumed Mean of Neutral Point*	*Mean*	*Mean as % of Maximum Possible*	*Average Item Rating*	*Standard Deviation*
1. Administrative Practices	9	18.0	15.580	58	1.73	7.164
2. Professional Work Load	9	18.0	19.603	73	2.18	5.017
3. Non-Professional Work Load	6	12.0	12.808	71	2.13	3.827
4. Materials and Equipment	8	16.0	15.887	66	1.99	5.768
5. Buildings and Facilities	7	14.0	14.432	69	2.06	4.506
6. Educational Effectiveness	10	20.0	20.280	68	2.03	6.439
7. Evaluation of Students	10	20.0	18.598	62	1.86	6.333
8. Special Services	8	16.0	15.420	64	1.93	5.153
9. School-Community Relations	7	14.0	13.088	62	1.87	5.142
10. Supervisory Relations	10	20.0	21.360	71	2.14	7.213
11. Colleague Relations	7	14.0	15.212	72	2.17	5.011
12. Voice in Educational Programs	8	16.0	13.234	55	1.65	5.217
13. Performance and Development	10	20.0	20.220	67	2.02	6.600
14. Financial Incentives	9	18.0	14.628	54	1.63	5.227

*Items on each dimension were scored 1 = disagree and 3 = agree.

tion of those groups or institutions that reflect different types or levels of the variables of interest must be undertaken with great care.

3. *Data Aggregation*: In most environmental studies student perception scores are pooled thereby masking individual differences in pupil perceptions. This is in conflict with some data (e.g., Bruner, 1958, pp. 85-94) which suggest that individual personality factors influence perceptions. It would be best *not* to have the environment perception scores correlate with any personality characteristics of the rater.

4. *Instrument Format*: It is generally desired that an environmental instrument contain a relatively few but uncorrelated scales. Lack of interscale correlations should aid interpretation or profiling of the scores because all existing data support the multi-dimensionality of environments. Many data reduction multivariate techniques are available and statisticians are becoming more sophisticated every day in their techniques for obscuring data. Beware of factor analysts bearing oblique axes. Some compromise between comprehensiveness and comprehensibility is desired.

5. *Data Analysis*: It is unlikely that instruments that are sensitive to rater personality would provide the desired high degree of between-institution/class variance and low within variance. We are usually looking for common environmental charactertistics not common personality characteristics unless students having a personality aspect in common have been grouped together or have been attracted to the same institutions.

6. *Convergent and Divergent Validity*: Evidence from a variety of sources bearing on validity should be sought: a multi-method approach. Like-named scales across measures should show some positive relationship. They should show some relationships with objective data. But it was also noted that for most applications, environmental scores should not correlate with characteristics of the inhabitants observers, or respondents. Monitoring with time-sampling techniques would also yield useful data.

The assessment of learning environments or climates is not an easy task. Due to the complexity of the phenomena involved, a variety of approaches should probably be tried. With the expenditure of time and effort (and if you are lucky some N.I.E.

monies) reliable and valid environmental measures can be developed and put to good use.

PROBLEMS IN THE DEVELOPMENT AND APPLICATION OF AFFECTIVE MEASURES FOR USE WITH INDIVIDUALS

There are a number of factors and influences which have inhibited the development and use of affective measures in our classrooms. Some of these are characteristic of teachers, others refer to technical problems.

Reluctance to Consider Affective Variables

Many teachers are reluctant to become involved with teaching and evaluating affective objectives. Some feel that these kinds of learning outcomes are of only minor importance or that this is an area where education has no province. Admittedly, affective objectives can prove to be a source of controversy, but the conscious avoidance of attacking affective objectives directly represents an affective stance by the teacher. Teachers are also rightly concerned about violating the students rights to privacy. It is becoming increasingly difficult to maintain a value free posture in contemporary society. By now educators must accept the fact that affective outcomes are equally as important as cognitive ones and therefore deserve equal treatment and time.

Another contribution to teacher reluctance rests on the educational philosophy of the local school board and administration. If those in control of the curriculum are reticent about the school becoming involved in the pursuit of affective objectives, obviously teachers will shy away from taking initiative on their own. How often have we heard a teacher remark that, "My principal wouldn't approve." Tyler (1973) has recently noted an allied political problem. He observes that the function of school in a democratic society is to help students gain the knowledge and skill necessary to increase independence in judgments and action, but *not* to indoctrinate particular political or sectarian views.

A frequently voiced complaint of the harried classroom teacher is that he or she does not have time to devote to developing and applying adequate assessment procedures. When a time constraint occurs, the first type of assessment to be short-changed is the affective. When cognitive and affective objectives are in competition for time, attention and resources, the affective tend to be the first to go, thereby contributing to the erosion of affective development in our students. It is a matter of priorities. Time must be found to consider these important variables. Some time may have to be taken from instructional activities because of the particularly intimate association between teaching for values and testing for them. But instruction can focus on both the so called cognitive and affective objectives simultaneously.

Lack of Faith

This problem area is difficult to overcome as it is anchored in one's belief or lack of belief that paper and pencil inventories and scales can measure classroom variables related to meaningful behavior. There is considerable evidence to support the contention that measured attitudes do relate to important school outcomes. It is almost a truism that if a student feels good about himself, he will probably feel good about learning.

Articifiality of Situation

To some extent this specific problem relates to the larger problem area of the validity of the instruments we use to measure. To ask an individual what he would do and then make the assumption that he would in fact do it, if the opportunity presented itself, is somewhat artificial. The fact that great reliance is placed on inference in assessing affective outcomes must be accepted. There is, admittedly, a real question of the relation of verbalized and actual behavior. All possible efforts must be made to help insure that the relationship is as strong as is possible with known techniques or the development of new ones. Does the apparent lack of realism mean that we should discount attitudes and values. No indeed! It does mean that great care needs to be exercised in assessing such variables. The fact that an individual is willing and does make a response to an inventory has some meaning. The less personally controversial or threatening the attitude being assessed the greater the likelihood of a valid response.

Public vs Private Attitudes

The most reasonable approach to the interpretation of self-report statements on affective measures is to accept them as public declarations rather than reflections of typical or private characteristics. Context plays an important part in determining the validity of self-reports. Attitudes toward the woman's liberation movement in general, auto styles, and men's clothing are relatively non-threatening general topics. When an individual is pressed to make specific revelations concerning his relationship with his wife's or daughter's liberation, his specific relations with minority group members, or his vote for a particular political candidate in the last election, the respondent will more likely attempt to conceal his true feelings. Many of the affective variables dealt with in the classroom setting, however, are of the less personally threatening variety and therefore more readily lend themselves to assessment.

Lack of Knowledge of Techniques

An obviously inhibiting factor in any assessment program is lack of knowledge about methods that can be employed in measuring affective outcomes. Most teacher training programs avoid any systematic attention to the specification or measurement of affective outcomes. Even test and measurement classes devote only minimal time to the topic. It is of little wonder then that teachers-in-the-field pay little attention to affective variables. Publications by Beatty (1969), Oppenheim (1966), Shaw and Wright (1967) and Mayhew (1965) have made significant contributions to technical literature illustrating the variety of measurement approaches available. Assessment in the classroom setting is not all that difficult; witness Corey's (1943) early efforts.

In addition to these five factors, there are a number of technical problems related to the development and application of affective measures that need to be considered. Such factors as semantics, fakeability, control of social desirability, and response sets and styles are just a few trouble spots in instrument development and validation.

There remains the perplexing problem of what to do with affective data. Obviously, they can be examined for individual and/or group progress. They can be used by the teacher to monitor classroom atmosphere. What about reporting such data. There are probably two compelling reasons for *not* treating affective data in the form of grades. First, it is difficult at best to prescribe, or proscribe within limits, specific affective outcomes in a democratic educational setting. Second, there is probably little hope of ascertaining the real feelings of students if rewards, sanctions, or penalties were attached or associated with any particular type of attitude.

If real efforts are to be directed toward the implementation and assessment of affective outcomes in our classes, significant changes in aspects of our teacher training programs at both inservice and pre-service levels will have to be instituted. Teachers must be provided with greater experience in methods useful in helping students examine their attitudes and values. In addition the ways in which our classrooms are organized will have to be restructured. In many locations this already is well underway. These requirements would seem particularly crucial due to the increased emphasis on individualized student learning programs and competency based teacher training programs.

APPROACHES TO THE ASSESSMENT OF AFFECTIVE OUTCOMES

Approaches to the assessment of affective variables are limited only by creativity and motivation. Many methods have been developed by psychologists and sociologists in their studies of human behavior, but few have been communicated to, or translated for use by, educators.

Cattell, Heist, and Stewart (1950), after extensive review of the literature and personal research, have identified numerous methods that can be applied in the assessment of attitudes and sentiments, or as they refer to them, "dynamic traits." Selections from their list and some additional methods follow:

1) *Money*. The amount of money an individual spends on certain activities and courses of action is a direct reflection of his attitude and interest.

2) *Time*. The amount of time an individual devotes to certain activities is, to some extent, a reflection of his attitude toward them.

3) *Verbal expressions*. A host of affective assessment methods use verbal expressions. Thurstone, Likert, Semantic Differential, and opinionnaire methods are illustrative.

4) *Measures of attention/distraction*. Records of length of time an individual attends to a stimulus, or a ranking of stimuli (e.g., pictorial) according to responsiveness to them, could profitably be used as measures of attitudes. Failure to respond to certain stimuli is also meaningful behavior.

5) *Fund of information*. The amount or type of information an individual possesses about a certain object, individual, or issue is to some extent a reflection of his attitude or interest.

6) *Speed of decision (reaction time)*. It may be that decisions are made more quickly about questions on which the subject has the strongest convictions.

7) *Written expressions (personal documents)*. Analysis of such documents as biographies, diaries, records, letters, autobiographies, journals, and compositions can be very revealing of an individual's attitudes. A *personal document* has been defined by Allport (1942) as any self-revealing record that intentionally or unintentionally yields information about the structure, dynamics, and functioning of the author's mental life.

8) *Sociometric measures*. Analysis of friendship choices, social distances, preferences and the general social structure of a classroom can be very informative about attitudes.

9) *Misperception/apperception methods*. Provided with ambiguous stimuli, an individual may be tempted to perceive them in accordance with his own interests, attitudes, and wishes.

10) *Activity level methods*. There are a number of measures of the individual's general excitement level in response to a stimulus, among them (a) fluency (amount written), (b) speed of reading, and (c) work-endurance.

11) *Observations*. The use of standardized reports systematically gathered by trained recorders operating within the limits of an explicitly stated frame of reference has provided extremely

valuable data on attitudes per se and on the operation of these attitudes within the individual (see Amidon and Hough, 1967).

12) *Specific performances and behaviors*. An individual's behavior can illustrate his attitudes and their influences. It is argued by some that behavioral measures are by far the most valid. The indirect methods we commonly use, however, can provide valid data if reasonable precautions are taken and stringent criteria are employed during the developmental stages. Webb, et al., (1966) have written an extremely valuable reference work with examples of behavioral measures and observational methods.

13) *Physiological measures*. The use of autonomic and metabolic measures can provide useful data in controlled situations. Psychogalvanic response, pulse rate, muscle tension and pressure, and metabolic rate are some of the procedures employed.

14) *Memory measures*. Instructing an individual to learn given material, varying the controversial nature of the content, introducing an unrelated activity to distract the subject, and then asking him to recall all or part of the original material is one approach to the use of memory as an instrument of attitude assessment. The selective operation of memory in reminiscence, dream, or fantasy may also be analyzed.

15) *Simulation*. The use of paper and pencil, visual, auditory, computer, or role-playing simulations can provide worthwhile data about inter-personal skills and attitudes (Levine and McGuire, 1970).

SUGGESTED AND ILLUSTRATIVE USES OF MEASURES OF AFFECT

Applications of affective measures fall into five general categories: 1) classroom application, 2) screening and selection, 3) counseling, 4) research, and 5) program evaluation.

Classroom Applications

The imaginative classroom teacher can create many situations in which the use of standardized or homemade affective measures makes a real contribution to the instructional program. The Work Values Inventory (WVI) (Super, 1970), for example, could be used on a pre- or post-basis to assess changes in work values associated with a unit on the "world of work," or the WVI itself could be used as a starting-point in exploring various occupations. It might be helpful to have students estimate their scores before taking the test, and then compare these estimates with the test results. Individual student scores or class means could then be compared with selected normative data. The study of vocations could be stimulated by this method; discussion might revolve around known differences between occupational groups.

Screening and Selection

The School Interest Inventory (Cottle, 1969) is an instrument that illustrates well the sensible use of an affective measure. The SII is used to identify potential dropouts. It is suggested that the SII be used on an intrainstitutional basis, so that a student's scores are compared only to those of other individuals in the same school. Students in the seventh or eighth grade could be administered the SII, and their scores ranked from highest to lowest within grade and sex. (Higher scores indicate a greater probability of the student's dropping out of school). Using any number of criteria, e.g. a cutoff score of 25 or above or selection of the top 20 percent, one could identify students who might benefit from counseling. Counselees could consider the possibility of continuing in the same or another course of study, or explore vocational and social adjustments that do not require a high school diploma. The counselor or teacher may also wish to set up "rap groups" in which personal, social, or vocational problems could be explored. Obviously, the use of a test as a screening instrument should be undertaken in conjunction with other relevant data. School achievement records, attendance, teacher's opinions, and age relative to school grade need to be considered.

Personnel managers frequently find that affective measures are useful in the hiring and placement of special classes of employees, and that scores may be related to job success. It is imperative when an affective measure is used in this manner that its relevance be demonstrable.

Counseling

Perhaps the major uses of standardized affective measures involve counseling. The value of such measures to stimulate a student to look at himself cannot be overestimated. The test can be used as a starting point to help establish rapport in the counseling interview. Asking the student to predict his scores and then comparing his prediction with the actual results can be beneficial. A diagnostic profile such as those that frequently accompany so many interest and attitude survey instruments is very useful in this kind of activity. Descriptions of the subscores can also be used as a basis for discussion. Student involvement in the actual task of profiling is recommended.

Research

There are numerous fields of research using affective measures that might prove of interest to the educator. The authors of the Study of Values, for example, note that it has been used to research the following topics:

1) Differences in scores of those in different college majors and occupational, religious, ethnic, and nationality groups.

2) Changes in values over time, and these changes as functions of specific training and educational experiences.

3) Relationships with other attitude-, interest-, and cognitive-style measures.

4) Relationships between friendship choice and sociometric status.

Such a list is really endless.

Program Evaluation

Another area in which affective measures are achieving great popularity is program evaluation. Curriculum evaluation is receiving increased attention from educational measurement and assessment experts and consultants. Most state and federal educational programs require the assessment of affective variables, and local school systems are also becoming conscious of these important outcomes. One interesting development along this line is the construction of the Minnesota School Affect

Assessment battery which is applied on a system-wide basis (Johnson, 1974). A combination of semantic differential concept rating scales, and rated statements make up the battery, suitable from K-12. Another extensive system has been developed by Bills (1975).

Measures of such variables as attitude toward school, respect for self, and appreciation of artistic efforts are illustrative of educational product and process outcomes in a comprehensive evaluation system. The self-concept is a personal attribute that is given considerable attention in many educational program evaluations.

ILLUSTRATIVE MEASURES OF AFFECT

It is impossible to adequately survey the kinds of measures of affect available from commercial and selected non-profit organizations, nor to review the tremendous variety of custom-made devices that can be developed. There are many sources, in addition to Buros' compendia, that might be consulted. Among the best references are those by Robinson and Shaver (1973), Frith and Narikawa (1972), Chun, Cobb and French (1975), Walker (1973), Bonjean, Hill and McLemore (1967), Lake, Miles and Earle (1973), Shaw and Wright (1967), Beatty (1969) and Payne (1974). Following are four instruments that reflect the flavor of instruments that are available or that can be produced. These instruments could be used on an individual or group basis.

School Sentiment Index

The SSI, available from the Instructional Objectives Exchange (Frith and Narikawa, 1972), measures several dimensions of attitudes toward school. This eighty item inventory requires students to indicate, anonymously, their perceptions of, or attitudes toward, various aspects of school (dichotomous response, True-False). The reliability, both internal consistency and stability, have been established on a variety of student groups. Modifications of the basic SSI structure are available to cover the entire grade range. Administration time is about 30 minutes. The major dimensions are presented in Table 7.

TABLE 7
Dimensions of the School Sentiment Index, Number of Items on Each Dimension, and Sample Items

Dimension	No. of Items	Sample Items (True-False)
Teacher Mode of Instruction	12	"My teachers make sure I always understand what they want me to do."
Teacher Authority and Control	12	"In my classes, the teachers allow us to make many decisions together."
Teacher Interpersonal Relationship with Pupils	12	"My teachers are interested in the things I do outside of school."
Learning	10	"The biggest reason I come to school is to learn."
Social Structure and Climate	10	"This school has events all the time that make me happy I attend school here."
Peer Relationships	12	"Other students bother me when I'm trying to do my school work."
General Attitude Toward School	12	"Each morning I look forward to coming to school."

Remmers' Generalized Standard Scales

Using the equal-appearing intervals method, H. H. Remmers and his colleagues have developed a series of "standard scales." A generalized standard attitude scale is one that can be applied to any of a selected class of objects. The scale may be used to measure attitudes toward any given subject, for example, by inserting in the appropriate space the name of the subject. The statements in the scale remain the same and have the same values regardless of the subject chosen. A sample master scale developed by Remmers and Silance (1934) is presented below. The scale values in parentheses following each item would not be included when the scale was duplicated for use.

A SCALE FOR MEASURING ATTITUDE TOWARD SCHOOL SUBJECTS

Form A

Directions: Following is a list of statements about school subjects. Place a plus sign (+) before each statement with which you agree, and a minus sign (–) before each statement with which you disagree with reference to each of the subjects listed at the left of the statements. Your score will in no way affect your grade in any course.

Science	English	Math	
			1. I have this subject. (0.6)
			2. This subject is the most undesirable subject taught. (0.7)
			3. I detest this subject. (0.8)
			4. I look forward to this subject with horror. (1.0)
			5. This subject is disliked by all students. (1.3)
			6. It is punishment for anybody to take this subject. (1.5)

Values: Ethical, Moral and Social

The forced-choice format has been widely adopted in the construction of inventories of affective variables. This method has many possibilities for use in classroom measurement. Two examples will illustrate possible applications. The first is from a scale, the Personal VEMS test (VEMS stands for Values: Ethical, Moral and Social), developed by Gardner and Thompson (1963) in their investigation of social values governing interpersonal relations among adolescent youth and their teachers. The VEMS requires a verbal response indicating the action that ought to be taken when confronting certain problem situations. Each decision implies the selection of one value over another. The values in question are loyalty, honesty, truthfulness, kindness, generosity, conformity and impunitiveness. In an effort to encourage the respondent to become ego-involved in the situation, the respondent is asked in some items to supply the name of his best friend as a participant in the problem situation. Following are two sample items:

> You have just taken an important true-false examination in English. Your teacher has asked you to exchange papers so that you can grade each other's papers as she reads the answers aloud. You exchange papers with your best friend ________________, who is seated near you. He slips you a note which reads, "Please change a few of my answers when they are incorrect. I have to get a passing mark on this test!"
>
> WHAT DO YOU THINK YOU OUGHT TO DO?
>
> ______A. Help your friend so that he will get a passing mark on the test.
> ______B. Mark his paper in the same way you would grade the paper of any other classmate.

In this item, alternative A is scored on the Loyalty scale, and B on the Honesty scale.

You and your classmate play a clarinet duet for the school assembly. There was much applause. Feeling rather pleased with her performance, your classmate says to a group of people you are standing with, "I guess I played it just about perfectly, didn't I?" You know that she squeaked a little on some of the high notes and that her timing was faulty in a number of instances.

WHAT DO YOU THINK YOU OUGHT TO DO?

______A. Be generous and say to the person next to you, "She certainly was terrific today."

______B. Say, "It was a good performance but not perfect. You'd better do some practicing on those high notes!"

On this item, alternative A is scored on the Generosity scale and B on the Truthfulness scale.

Preferred Job Characteristics Scale

Farquhar and Payne (1963) have described the development of an instrument aimed at assessing relative preferences for statements correlated with occupational motivation. Beginning with a set of eight alternatives describing high achievement motivation and eight describing low achievement motivation, they constructed a 64-item pair scale by combining high and low alternatives. Two sample items from the scale follow:

I prefer:
  1. A job where my opinion is valued, or
  2. A job with short working hours

I prefer:
  1. A job which does not tie me down, or
  2. A job where I could decide how the work is to be done

In the first item, alternative 1 is the high-motivation alternative; in the second, it is alternative 2.

Self-Concept Measures

One of the most important dimensions of the human organism has been labeled self-concept by psychologists. Shavelson, Hubner, and Stanton (1976) have recently noted the viability of this construct, as well as pointed out the needed kinds of research. In support of the construct, evidence suggests that self-concept is: 1) organized and structured, 2) multi-faceted, 3) generally stable, 4) hierarchical, 5) developmental, 6) evaluative and differentiable from other constructs. Examples of representative devices are the Michigan State Self-Concept Scale, Coopersmith's Self-Esteem Inventory, Ira Gordon's How I See Myself, the Piers-Harris' The Way I Feel About Myself, and Paulene Sears' The Self-Concept Inventory.

IN SEARCH OF REMEDIES

What are some steps that educators and behavioral scientists might take to remedy the problem situations surrounding the assessment of affective educational outcomes and learning environments?

With the apparent shift in funding posture in Washington, particularly at N.I.E., researchers should now have the opportunity to initiate from the field proposals to investigate a great variety of topics. There are both *basic* and *applied* problems in need of investigation. Despite the presence of a substantial amount of theoretical literature, very little is really known about the composition and development of affective variables in an educational setting. What are the contributors to, and modifiers of, behaviors related to the expression of attitudes, interests, and values? We need basic data. In looking at the applied area, one can see many psychometric questions. Such questions as which measurement techniques would optimize the assessment of which kinds of affective objectives for what kinds of populations need to be addressed. We need more studies utilizing the Campbell-Fiske methodology called the multi-trait, multi-method matrix.

Another approach or thrust, and this may sound silly to many, would be to employ the sensitivity methods that many minority and psychological movements have used successfully to bring about increased awareness of issues. In teacher training institutions such consciousness-raising might be approached both formally, through changes in curricular emphasis, and informally, through the use of grouping and organizational development techniques. That old course in tests and measurements could stand to be over-hauled so as to include a greater emphasis in the assessment of affective outcomes.

In addition, instructional supervisors, lead teachers, curriculum directors, and school administrators can do a number of things to assist classroom teachers in overcoming some of the obstacles to effective assessment of affective learning outcomes. Some of these strategies are as follows:

1) Work with teachers in generating general curriculum designs which will be responsive to affective needs of students.

2) Assist teachers in the writing of affective objectives so that there will be at least some conscious awareness that these outcomes must be considered.

3) Help develop instructional experiences for students where both cognitive and affective outcomes are integrated.

Our professional organizations need to push for greater recognition of the importance of affective variables. There exists a Special Interest Group in the American Educational Research Association that deals with Affective Education. In addition a Task Force on Affective Measurement has been activated within the National Council on Measurement in Education. More and similar organizations need to focus some portion of their energy and resources on this vital area of education.

Suffice it to say that the affective components of education are important and that they should be and can be assessed. It is, as Stake and Gooler (1971) have noted, a matter of priorities and effort.

In closing let me relate a story told about young Peter Rabbit:

> It seems that Peter Rabbit was a lively little boy bunny who had the bad habit of stroking the little girl bunnies on the whiskers at every opportunity. His mother, wanting desperately to cure her son of this annoying habit, said to him, "Peter Rabbit, if you don't stop this, you are going to turn into a goon." Well, that startled Peter and he did stop — for three weeks. But one moonlit night he hopped out onto a

> clover field in which several little girl bunnies were playing and sidled up to one and stroked her on the whiskers.

And, sure enough, he turned into a goon! Now the moral of this story, I'm sure, is obvious: "Hare today and Goon Tomorrow."

So it is with many of the answers of our affective measurement problems. If we don't start to search now for some of them, tomorrow may be too late.

REFERENCES

Aiken, Jr., L.R. Attitudes toward mathematics. *Review of Educational Research*, 1970, *40*, 4, 551-596.

Allport, G. W. *The use of personal documents in psychological science*. N.Y.: Social Science Research, 1942.

Amidon, E. J. and Hough, J. B. (eds) *Interaction analysis: Theory, research, and application*. Reading, MA: Addison Wesley, 1967.

Anderson G. J. Effects of classroom social climate on individual learning. *American Educational Research Journal*, 1970, 7, 2, 135-152.

Anderson G. J. *The assessment of learning environments: A manual for the Learning Environment Inventory and the My Class Inventory*. Halifax, Nova Scotia, Canada: Atlantic Institute of Education, September, 1973.

Anderson G. J. & Walberg, H. J. Classroom climate and group learning. *International Journal of Educational Sciences*, 1968, *2*, 175-180.

Anderson, G. J. and Walberg, H. J. Learning environments. In H. J. Walberg (ed.) *Evaluating educational performance*. Berkeley, CA: McCutchan, 1974.

Astin, A. W. The Inventory of College Activities (ICA): Assessing the college environment through observable events. Paper presented at the annual meeting of the American Psychological Association, Chicago, 1965.

Astin, A. W. and Holland, J. L. The environmental assessment technique: A way to measure college environments. *Journal of Educational Psychology*, 1961, *52*, 308-316.

Barker, R. G. *Ecological psychology: Concepts and methods for studying the environment of human behavior*. Stanford, CA: Stanford University Press, 1968.

Bassham, H., Murphy, M. and Murphy, K. Attitude and achievement in arithmetic. *The Arithmetic Teacher*, 1964, *11*, 66-72.

Beatty, W. H. (ed.) *Improving educational assessment and an inventory of measures of affective behavior*. Washington, D.C.: Association for Supervision and Curriculum Development, 1969.

Bills, R. E. *A system for assessing affectivity*. University, Alabama: University of Alabama Press, 1975.

Bonjean, C. M., Hill, R. J., and McLemore, S.D. *Sociological measurement: An inventory of scales and indices*. San Francisco: Chandler, 1967.

Bruner, J. S. Social psychology and perception. In E. E. Maccoby, T. R. Newcomb and E. L. Hartley (eds.) *Reading in social psychology*. (3rd Edition). N.Y.: Holt, Rinehart, and Winston, 1958.

Cattell, R. B., Heist, A. B., and Stewart, R. G. The objective measurement of dynamic traits. *Educational and Psychological Measurement*, 1950, *10*, 224-248.

Cayne, R. E. Relationship of teacher personality to classroom social climate. *AERA Abstracts*, 1970, 40.

Chun, K., Cobb, S., and French, Jr., J. R. P. *Measures for psychological assessment (A guide to 3,000 original sources and their application)*. Ann Arbor, MI: Survey Research Center, Institute for Social Research. University of Michigan, 1975.

Corey, S. M. Measuring attitudes in the classroom. *Elementary School Journal*, 1943, *43*, 457-461.

Coughlan, R. J. Dimensions of teacher morale. *American Educational Research Journal*, 1970, *7*, 221-235.

Coughlan, R. J. and Cooke, R. A. Work attitudes. In H. J. Walberg (ed.) *Evaluating Educational Performance*. Berkeley, CA: McCutchan Pub. Corp., 1974.

Domino, G. Interaction effects of achievement orientation and teaching style on academic achievement. *Journal of Educational Psychology*, 1971, *62*, 427-431.

English, H. B. and English, A. C. *Comprehensive dictionary of psychological and psychoanalytical terms*. N.Y.: Longsman, Green, 1958.

Farquhar, W. W. and Payne, D. A. Factors in the academic-occupational motivations of eleventh grade under-and over-achievers. *Personnel and Guidance Journal*, 1963, *42*, 245-251.

Frith, S. and Narikawa, D. (eds.) *Attitude toward school grades K–12*. Los Angeles, CA: Instructional Objectives Exchange, 1972.

Gagné, R. M. *The conditions of learning*.N.Y.: Holt, Rinehart and Winston, Inc., 1965.

Gardner, E. F., and Thompson, G. G. Investigation and measurement of the social values governing interpersonal relations among adolescent youth and their teachers. U.S. Office of Education Cooperative Research Project 259A (8418), 1963. Syracuse University (Mimeograph).

Getzels, J. W. and Thelen, H. A. The classroom as a unique social system. In N. B. Henry (ed.) *The dynamics of instructional groups. 59th Yearbook of the National Society for the Study of Education, Part II*. Chicago: University of Chicago Press, 1960.

Gordon, I. J. Affect and cognition: A reciprocal relationship. *Educational Leadership*, 1970, *27*, 661-664.

Heath, D. H. Affective education: Aesthetics and discipline. *School Review*, 1972, *80*, 353-371.

Hoepfner, R., et al. *CSE-RBS Test evaluation: Tests of higher-order cognitive, affective, and interpersonal skills*. Los Angeles: Center for the Study of Evaluation, 1972.

Johnson, D. W. Affective outcomes. In H. J. Walberg (ed.) *Evaluating educational performance*. Berkeley, CA: McCutchan, 1974.

Kahl, J. A. Some measurements of achievement orientation. *American Journal of Sociology*, 1965, *4*, 669-681.

Khan, S. B. and Weiss, J. The teaching of affective responses. In R. S. W. Travers (ed.) *Second handbook of research in teaching*. Chicago: Rand McNally, 1973.

Krathwohl, D., Bloom, B. and Masia, B. *Taxonomy of educational objectives. Handbook II: The affective domain*. N.Y.: David McKay, 1964.

Lake D. G., Miles, M. B., and Earle, Jr., R. B. (eds.) *Measuring human behavior*. N.Y.: Teachers College Press, 1973.

Levine, A. G. and McGuire, C. H. The use of role-playing to evaluate affective skills in medicine. *Journal of Medical Education*, 1970, *45*, 700-705.

Lewin, K. *Principles of topological psychology*. N.Y.: McGraw-Hill, 1936.

Maguire, T. O., Goetz, E. and Manos, J. Evaluation activities of the IPI Project Year 2: The evaluation of two instruments for assessing classroom climate for primary grades. Alberta Human Resources Research Council, 1972.

Mayhew, L. B. Measurement of noncognitive objectives in the social sciences. In H. D. Berg (ed.) *Evaluation in social studies* *thirty-fifth yearbook*. Washington, D.C.: National Council on Social Studies, 1965.

Menne, J. W. Techniques for evaluating the college environment. *Journal of Educational Research*, 1967, *4*, 219-225.

Murray, H. A. *Exploration in personaltiy*. N.Y.: Oxford University Press, 1938.

Nunally, J. C. *Psychometric theory*. N.Y.: McGraw-Hill Book Co., 1967.

Oppenheim, A. M. *Questionnaire design and attitude measurement*. N.Y.: Basic Books, 1966.

Pace, C. R. and Stern, G. G. An approach to the measurement of psychological characteristics of college environments. *Journal of Educational Psychology*, 1958, 59, 269-277.

Payne, D. A. *The assessment of learning: Cognitive and affective*. Lexington, MA: D.C. Heath, 1974.

Payne, D. A., Ellett, C. and Klien, A. *The verification and validation of principal competencies and performance indicators, V. 1* Report to the Georgia State Department of Education. Athens, University of Georgia, 1975 (Mimeograph).

Perkins, M. L. *A canonical correlation analysis of the relationships among school climate, teacher, morale, and educationally-relevant performance of fourth grade students*. Unpublished doctoral dissertation, University of Georgia, 1976.

Pervin, L. A. A twenty-college study of student x college interaction using TAPE (Transactional Analysis of Personality and Environment): Rationale, reliability, and validity. *Journal of Educational Psychology*, 1967, *58*, 290-302.

Randhawa, B. S. and Fu, L. L. W. Assessment and effect of some classroom environment variables. *Review of Educational Research*, 1973, *43*, 303-321.

Remmers, H. H. and Silance, E. B. Generalized attitude scales. *Journal of Social Psychology*, 1934, *5*, 298-312.

Ripple, R. E. Affective factors influence classroom learning. *Educational Leadership*, 1965, *22*, 476-480.

Robinson, J. P. and Shaver, P. R. *Measures of social psychological attitudes*. Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI, 1973.

Scriven, M. Student values as educational objectives. In *Proceedings of the 1965 invitational Conference on Testing Problems*. Princeton, NJ: Educational Testing Service, 1966.

Shavelson, R. J., Hubner, J. J., and Stanton, G. C. Self-concept: Validation of construct interpretations. *Review of Educational Research*, 1976, *46*, 407-441.

Shaw, M. E. and Wright, S. M. (eds.) *Scales for the measurement of attitudes*. New York: McGraw-Hill, 1967.

Shoben, E. J. Jr. Potency in the schools. *Teachers College Record*, 1962, *63*, 548-550.

Simon, A. and Bouer, E. H. *Mirrors for behavior: An anthology of observation instruments*. Philadelphia: Research for Better Schools, 1970.

Stake, R. E. and Gooler, D. Measuring educational priorities. *Educational Technology*, 1971, *11*, 44-48.

Steele, J. M., House, E. R., and Kerins, T. An instrument for assessing instructional climate through low-inference student judgments. *American Educational Research Journal*, 1971, *8*, 447-465.

Stern, G. G. *People in context: Measuring person-environment congruence in education and industry*. N.Y.: John Wiley and Sons, Inc., 1970.

Super, D. E. *Work values inventory*. Boston: Houghton Mifflin, 1970.

Tyler, R. W. Assessing educational achievement in the affective domain. *Measurement in Education*, 1973, *4*, 1-3.

Walberg, H. J. The social environment as a mediator of classroom learning. *Journal of Educational Psychology*, 1969, *60*, 443-448.

Walberg, H. J. and Ahlgren, A. Prediction of the social environment of learning. *American Educational Research Journal*, 1970, 7, 153-167.

Walberg, H. J. and Anderson, G. J. Classroom climate and individual learning. *Journal of Educational Psychology*, 1968, *59*, 414-419.

Walberg, H. J., Sorenson, J. and Fischback, T. Ecological correlates of ambience in the learning environment. *American Educational Research Journal*, 1972, Vol. 9 139-148.

Walberg, H. J. and Welch, W. W. A new use of randomization in experimental curriculum evaluation. *School Review*, 1967, *75*, 360-377.

Walker, D. K. *Socioemotional measures for preschool and kindergarten children: A handbook*. San Francisco: Jossey-Bass, 1973.

Walsh, W. B. *Theories of person-environment interaction: Implications for the college student*. Iowa City, IA: American College Testing Program, 1973.

Webb, E. J., Campbell, D. T., Schwartz, R. D. and Sechrest, L. *Unobtrusive measures: Nonreactive research in the social sciences*. Chicago: Rand McNally, 1966.

NSPER: 76

Phoenix Presentation

Dr. Walcott H. Beatty

AFFECTIVE MEASUREMENT AND THE SELF CONCEPT

Walcott H. Beatty
San Francisco State University

> In recent years increasing attention has been given to the importance of feelings of children in the classroom. The earlier notion of training the intellect has given way to the belief that education must deal with the whole child. (Beatty, 1952)

I feel that I have come full circle. That was the opening paragraph of my Ph.D. dissertation at the University of Chicago, written twenty-six years ago.

The dissertation was a validation study of a way of using classroom observations for determining the climate of feelings in a classroom. Looking back on those words, I think I was wrong. Very little attention was being paid generally to the feelings of children in the classroom. At the University, Herb Thelen had set up a laboratory classroom in which it was possible to study the many interactions which take place. Two of his students, John Withall (1951) and Ned Flanders used these facilities and made significant contributions to the development of observation techniques. In fact, it is my belief that Flanders' (1965) study sponsored by the U.S. Office of Education entitled, "Teacher Influence, Pupil Attitudes, and Achievement," was one of the great breakthroughs in educational research.

My early optimism about growing concern for the children's feelings came largely from reading a book called *Emotion and the Educative Process* by Daniel Prescott (1938). It appeared obvious that the connection between emotions and learning could no longer be ignored. At about the same time the second symposium on *Feelings and Emotions* was held at the University of Chicago. It was here that I heard a paper read by Carl Rogers (1950) which influenced me greatly. Much of what I will be discussing had its beginning there.

Despite the rather slow progress in this area with the passage of time, I still feel optimistic. It is indeed quite possible that this conference, sponsored by Phi Delta Kappa, may lead to follow up work by those attending which could well be a milestone in the study of feelings and emotions. I believe progress has been held up partly because measurement in this area has been so difficult.

One of the most complete systems for "assessing affectivity," developed by Bob Bills (1975), leaves key questions undiscussed. Two of these are: what is the role of the school in developing affective behavior, and what implications does a completed assessment have for steps to be taken by the school? I state this, not as criticism, but as an indication of where we stand at the present time, and therefore, how important this conference is. In my opinion, Bills has gone farther than anyone else up until now, to make it possible to answer these questions. I shall be discussing his work later.

In formulating my presentation I have been guided by Ralph Tyler's (1969) concluding remarks in the 68th Yearbook of the National Society for the Study of Education on educational evaluation. He states that the foundations upon which future developments in educational evaluation will take place are four essential and interrelated operations:

1) Clarifying the particular educational function for which evaluation is needed. In this case, it will be the role of affect in learning.

2) Formulating a body of theory, including concepts and assumptions relevant to this educational function. This will, I hope, be my main contribution.

3) Selecting and developing evaluation instruments and procedures that are consistent with this body of theory. I will attempt to give samples of these.

4) Revising on a continuous basis, in the light of the information obtained by this process.

Following these steps, I will start by trying to clarify the meaning and importance of affect in the learning process. Then I will present a theoretical formulation based on a self concept theory of motivation and learning. Finally, I will be looking at measuring instruments and procedures which are consistent with the theory. Tyler's final operation, that of revising on a continuous basis, is something I hope we can all work on in the future.

AFFECT IN THE LEARNING PROCESS

First, consideration must be given to the problem of what is meant by "affect" and "affective behavior." We all experience feelings and emotions, but our experience is often not directly observable by another and it is not an easy concept to define. For my purposes today, I wish to stay with a very simple idea, that affect is the experiencing of positive or negative feelings, the awareness of pleasantness or unpleasantness. Furthermore, this experience has motivational properties. One moves toward or attempts to continue pleasant feelings; one seeks to move away from or to stop unpleasant feelings. All motivation has an affective quality and all feelings are motivating.

"Affective behavior" is a term used to refer to actions which we have learned are expressions of feelings or emotions, or which are clues to an underlying affect. Facial expressions, body posture, tone of voice, vigor of response, as well as some behaviors such as boasting, quarreling, reaching out, and others, are clues which our past experience has validated as having important affective meaning.

Words such as interests, attitudes and values are used to indicate some affective disposition. These are indicators that some cognitive matter has motivational affects attached to it. Distinctions can be made with regard to the extent that each of these is integrated into the total personality. *The Taxonomy of Educational Objectives: Affective Domain* (Krathwohl, et al., 1964) uses the continuum of internalization. As one proceeds from the lowest to the highest categories of responses, there is increasing emotional commitment, increasing complexity, increasing abstractness of the category, and increasing importance in the structure of the personality. These are useful ways of looking at affective elements, but actually in human behavior, these are really cognitive-affective responses. All cognitions have some affective character, and conversely, all affects have cognitive elements. The relationship is that of figure to ground. Either the cognitive or the affective component can be in figure. Typically, we tend to stress one or the other. Mathematical ideas are cognitive, but within the awareness of an individual, they always carry affective meaning as well. One is motivated toward or away from such ideas with differing degrees of intensity. Thus, elements of the cognitive realm which we are able to specify can be used as a basis for knowing that there is an affective quality which can also be investigated. Most of the studies in this area have been with regard to attitudes toward subject matter.

In the same way that we have feelings about cognitive matters, we have cognitions about affective matters. We know the behavior which is associated with various feelings. This underlies skill in dramatic acting and, at the same time, is what enables us to infer how others are feeling.

We are born with the mechanisms which give rise to feelings and emotions. There may be some unlearned affective reactions, such as those to pain, but it appears that for the most part, we learn to associate positive or negative affect with our experiences. Thus, we learn to like foods and to dislike people on the basis of skin color or their religious views. This process, which Freud called cathexis, takes place all through life, and it takes place in the classroom. Because of the motivational nature of these affective attachments, it becomes important to measure these attachments and, perhaps, even more important to understand how they develop.

A distinction needs to be made between feelings and emotions. They are both present in the classroom and have important consequences for learning, but I would argue that for measurement purposes we are mainly concerned with feelings. The diagram below will illustrate this.

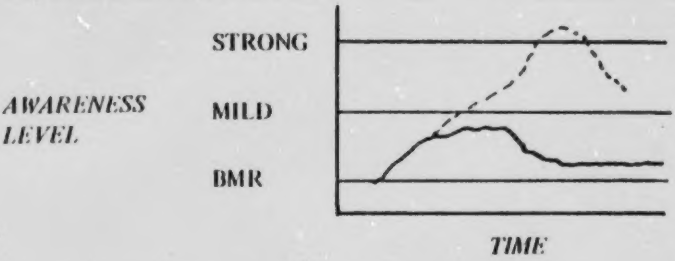


Figure 1. Energy mobilized in the body as indicated by the strength of emotion.

The lowest level of affect mobilizes the least amount of energy or motivation toward activity. This is roughly equivalent to the basal metabolic rate, which is just sufficient energy to maintain life processes. Usually, in a waking state, we would be at a somewhat higher level of energy mobilization, which I have labeled the "awareness level." We are aware of pleasant and unpleasant sensations associated with sensory inputs. If the sensory input calls for action, normally the energy level and strength of feeling increase. Such an increase starts activity to use the energy and respond to the feeling. In the diagram the arrow represents such an input. Perhaps in the classroom the teacher has asked the student a question. The upward curving line indicates that the question has released energy. The student would then answer the question, thus making a constructive use of the energy. When he has finished his answer the feeling-energy level may or may not return to a lower level. If the response was satisfactory, then the energy level tends to decrease. If, however, the student does not have a satisfactory response, then the energy level may continue to climb as I have indicated with the dotted line. The increased feeling climbs to the level of strong emotion. The student begins to squirm and feel uncomfortable. So much mobilized energy with no outlet can be contained only so long. If the student has no skills or knowledge with which to handle the demands of the situation, he must do something to get rid of the overwhelming feeling. This is the point at which he will run from the room, or yell, or hit someone, or cry; he takes any action which will release the energy and allow him to return to a more comfortable level. As observers, we would be well aware that the student was experiencing strong feelings. He would also have cognitions, but we would be less aware of them. The cognition in this case might be the knowledge that the instructor was the cause of all of his discomfort, and so, it is quite sensible to yell at the teacher or run away. The concern of the educator is to help the student learn to make adequate responses which enables the energy to flow productively and not evoke strong emotions.

Behavior at the level of strong emotion is indicative of the fact that the student is under too much pressure. Possibly this is because of the demands of the classroom, or it may be due to factors outside the school. Recurring behavior, indicative of unmanageable stress, suggests the need for special psychological help.

The above diagram illustrates short term fluctuations of feelings in response to immediate classroom demands. If the student develops persisting positive feelings toward learning, toward specified cognitive material such as literature or reading, and toward those values which make for effective personal functioning, then he will rarely feel stress and will feel successful in school.

In 1965, the journal *Educational Leadership* devoted an entire issue to affective learning. In the lead editorial, Early Kelley said: "It could well come about that this is one of the most important issues in the history of this publication. The reason for this statement is that it has now become abundantly clear from research and from reason, that *how a person feels is more important than what he knows*. This seems true because how one feels controls behavior, while what one knows does not. What one knows is used in behavior, to be sure, but the way it is used depends upon positive or negative feelings. It is possible to be a saint or a demon with similar knowledge" (1965, p. 455).

In 1969, Piaget, who has become one of our most respected researchers in developmental psychology, published his book, *The Psychology of the Child*. He stated: "There is no behavior pattern, however intellectual, which does not involve affective factors as motives . . . . Behavior is therefore of a piece. . . .The two aspects, affective and cognitive, are at the same time inseparable and irreducible" (1969, p. 158).

In research, the interest and findings are becoming more directly of concern to education. Ira Gordon (1970) reports studies in which the self concept of an individual, in relation to school, was the best single predictor of achievement in school. Perhaps the most comprehensive series of studies concerning the effects of affective factors on school performance are those of Bob Soar (1969, 1972). In a chapter in the 1970 ASCD Yearbook, he summarizes his own research and that of more than a dozen other researchers. The conclusions supported by all of this work clearly indicate that a more indirect, more open, more supportive style of teacher behavior increases the subject matter growth of pupils. In addition, such teacher behavior is associated with more favorable pupil attitudes toward school and increased growth in creativity. Furthermore, the effects of such a facilitative classroom climate continues after the pupil leaves the classroom.

In a study by Davis (1976), it was found that, in college level classes where the teacher used student ideas and student input was positively reinforced, the students reported positive attitude changes toward the subject matter. Fox, Lippitt, and Schmuck (1964) carried out research in an elementary school where they found that pupils who perceive their parents as supportive of the school utilize their abilities more fully. The same thing was true of pupil liking of the teacher. When they liked her, they made greater use of their abilities. If they were dissatisfied with the teacher, they were dissatisfied with themselves. In relationship to their peers, students made greater use of their abilities if they were well liked. The opposite was true for students who had few friends. More children made use of their abilities when more of the children were well liked.

This is only a small sample of the research in this area, and clearly, it supports the idea that a person's feelings are important in the learning process.

A SELF CONCEPT THEORY OF MOTIVATION AND LEARNING

I would like to proceed with Tyler's second operation and attempt to imbed these ideas about feelings in a model of learning and motivation. It seems possible that one of the reasons that progress in understanding and measuring feelings in the classroom is that it has been done on an ad hoc basis. Attitudes, for example, are measured without any clear idea of how the attitudes developed or what can be done if one wishes to change attitudes. I believe that this is because feelings and emotions can only be understood in relationship to the personality of the person experiencing the emotion. It is as a person is evaluating the meaning of an experience for his self that feelings arise (Beatty, 1969, p. 81). Furthermore, such a model has implications for what we should be measuring in the affective domain and for the kinds of instruments we should be using.

Perceived Self in the World

This is a self concept model based on the hypothesis that an individual organizes his experiences over time around a picture of himself or herself functioning in the world. This perception of one's self in the world develops as one is reacted to by others. The parents respond to the child as male or female. Each of these concepts is extremely complex. It includes how one dresses, what one plays with, who one plays with, and a whole constellation of behaviors associated with maleness or femaleness. The child is also responded to with love or rejection and he sees himself as he is seen, i.e., if he is loved, he sees himself as lovable.

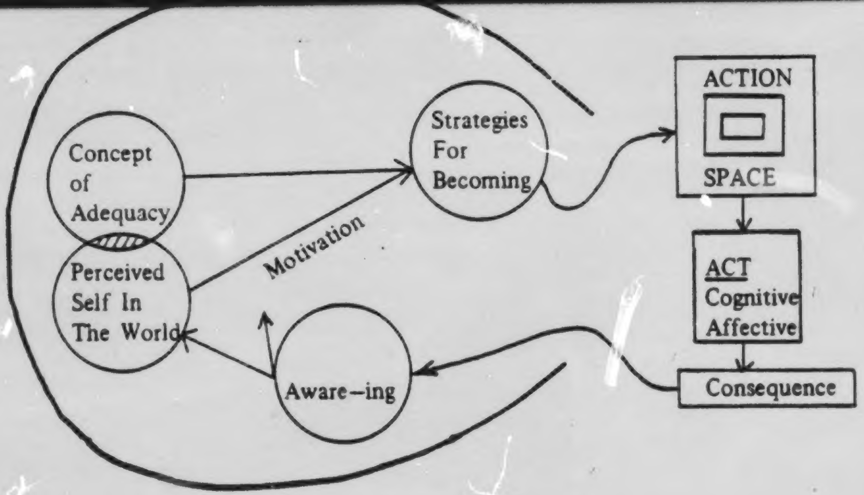
The child's behavior is also appraised as a member of a culture. The culture prescribes a way of life ranging from what is appropriate for one to eat to the most deeply held values about how one should lead one's life. This might be broken down into many more detailed parts, but the important point is the fact that the child experiences many, many actions and reactions which guide him in defining himself. The part of the self which emerges through this process, I call the perceived self in the world. To the child, this perception is who he is and it determines how he relates to other people in the world. In the early years, the child usually lives in a family that continues to respond relatively consistently and to reinforce these developing images. The images, therefore, become his *own* picture of himself. In time, the child learns about alternate ways to be and behave, but since these are incompatible with his perceived self, they are resisted and the self concept which he has developed becomes reality; that is the way he is. The perceived self changes as behaviors which he formerly identified as himself cease to be reinforced, and some new image, which *is* being reinforced, replaces it. This happens as the child moves from home to school and to other parts of the world, beyond his early experiencing. Thus, a child continues to change and grow gradually as his experience of himself in the world reinforces or denies expanded or additional conceptions. However, at any one time, his behavior will be consistent with the picture he holds of himself.

The Concept of Adequacy

At the same time that the perceived self in the world is developing, another part of the self concept is emerging. Not only does the child experience appraisals of what he is like, he is also appraised in terms of what he could or should be like. More important than this, he has the models of mother and father who are perceived as so much more powerful than the child. To him, as a very young child, they appear to be able to do everything. If only he could be like "Mommy" or "Daddy," then he too would be able to control the world and his own life. He conceptualizes that which he is able to understand, or of which he may be aware, as their qualities and tries to become like this. He imitates parental behaviors. Through this identification process he develops an additional part of the self concept which I call the Concept of Adequacy. It is a model of what one should be like if one is really going to be able to satisfy one's needs and to function effectively. This part of the self also grows and changes as the child grows. Later on, his mother and father are seen as having limitations and behaviors which cause negative affect in the child, and he supplements his concept of adequacy with additional identifications from other sources. At some point, it may be the Bionic Man, or a supportive teacher, or a star athlete. A person's picture of adequacy is a complex mixture of many identifications and is, before long, completely unique to that person. The concept of adequacy, like the perceived self, is reality to the individual but is, of course, unlike anyone else's reality.

Figure 2 is a diagram of this model of the self concept. As can be seen, the perceived self and the concept of adequacy overlap to some degree. As he grows, a person sees himself as adequate in some ways. In the general way in which he lives his life, the way he dresses and eats, his way of earning a living, all may seem adequate and the person has little desire to change them. However, a large portion of these two parts of the self concept do not overlap. A person's way of seeing himself is discrepant from his picture of adequacy. It is this discrepancy which is the major source of the motivation to learn and grow. Each person is continually striving to become more like his concept of an adequate self.

Since behavior is consistent with the self concept, and since a person is motivated toward adequacy, it is possible through the observation of behavior to infer characteristics of a person's self concept. This self concept consists of so many thousands of perceptions so that a full understanding would take much time and observation. Fortunately, the behavior which is an expression of the self concept can be organized around four nodal points. In


Concept
of
Adequacy
Perceived
Self In
The World
Motivation
Strategies
For
Becoming
ACTION
SPACE
ACT
Cognitive
Affective
Consequence
Aware–ing

Organizing Centers of the Self Concept (Feelings/Cognitions)

- Personal Worth
- Ability to Cope
- Ability to Express
- Autonomy (Ability to make choices)

Figure 2. A Self Concept Theory of Motivation and Learning (Beatty & Clark, 1968)

other words, I think all behavior is related to one or more of four organizing centers in the self concept.

Worth

The first of these organizing centers is the feelings and cognitions of personal worth. Each person has feelings about his perceived self related to how worthy he is. He also has a concept of what perfect worthiness would be like. He is motivated to learn ways in which he can behave so that it will increase his feelings of personal worth. Feelings of worth come from being loved. A child's experience of being wanted and included, of having his needs taken care of merely because he exists, builds a feeling of having worth. There are also times when needs other than his are given priority by his parents and he is not included. This means that he feels less worthy than it is possible to become. If this discrepancy between how worthy he perceives himself to be and his perception of an ideal state of worthiness is small, he is weakly motivated to try to become more included and valued, and to others, he appears secure. If this discrepancy is great, he appears jealous or resentful of others, and strives continually for attention.

Coping

A second organizing center in the self concept is around the feelings and cognitions concerned with coping in the world. The increasing competence that a child experiences as he grows, develops coordination, learns skills and acquires useful knowledge gives him increasing confidence that he is able to handle the demands of the world. The abilities of others, adults in general, expert athletes, famous people, and people with talents which are admired, set a standard of coping ability beyond that of the child. However, as long as he is progressing in a satisfactory way in his learning, he maintains confidence that he too will become an expert at coping. Learning to cope in our society is so complex that it requires schools and many years of work to become capable. If the demands of parents or of the school set unrealistic standards, or if their judgments of him find fault with him, he will lose confidence. Again, if the discrepancy between his perceived ability to cope and his concept of adequacy is small, he will feel challenged and confident as he learns. If the discrepancy begins to increase he will dislike school and want to escape. He will have negative feelings about learning.

The feelings of personal worth and of confidence that one can cope with one's world are two separate organizing centers. One can be loved even though he can not cope, and one can cope even if he is not loved. However, for the average person, they become considerably entangled. The child may have started out with unconditional love as a helpless baby. Later parents may withhold love when a child behaves "badly" or they may use love as a reward for "good" behavior. This tends to teach the child that one achieves worth through doing as others wish, that coping effectively is the secret of being prized. It is my belief that the schools are, to a large extent, responsible for this confusion in the child. The fact that much of what we insist that children learn in school leaves them unmotivated, pushes us to use approval, and threats to withhold it, as a source of motivation. The skill or knowledge should be giving the child a feeling that he is increasing his ability to cope. When he learns these things in order to get love it turns our evaluation system into a kind of global evaluation of the child. Since he is working for approval, an "A" means he is really loved, and an "F" is strong rejection. However, there is no way in which more knowledge or skill can bring feelings of worth for more than a transitory period. At the moment of great achievement, one may feel a glow of being admired, but it quickly fades and the child is still no more loved than he was before the achievement. Using love or approval as a motivating device may lead to greater productivity but it comes at a high psychic cost.

It is possible to disentangle these two sets of behaviors, so that an individual is motivated to learn things because he can clearly see that they enable him to cope better. At the same time, he can learn that worth comes from relating to people around him so that they care about him, merely because he is himself and not because he can cope well. I have hope that progress in measuring affect will contribute to a recognition of classroom processes which keep these two centers of motivation separate.

Expressing

I have said earlier that all experience is accompanied by a pleasant or unpleasant affective tone. The nature of the organism is such that a number of activities, completely unconnected with coping, are pleasant, and we are motivated to participate in them. Music, rhythm, color and form are examples. They bring feelings of fulfillment. It is my belief that our frenetic culture, with its emphasis on coping, has almost destroyed the arts. In the elementary grades, schools encourage expressive experiences: painting, sculpting, listening to music, marching and dancing, and discovering the beauties of nature. These activities tend to disappear by mid-elementary school and return later only as electives. It is this feeling of the need to express which I see as the third organizing center for behavior.

This kind of expressing brings feelings of delight or beauty, but they are not the only parts of the self which need to be expressed. The blocking effect of anger or fear, when we are unable to think or act effectively, is known to all of us. Strong positive emotions, such as great joy or love, distort our reactions if they are suppressed. The expression, "I'm so happy I think I'm going to bust," is indicative of the pressure such emotions have for expression. Unfortunately, our culture does not look favorably upon strong emotions. In the growing up process, most of us have experienced times when we were sent to our rooms to cry or were forbidden to express openly our feelings of anger. This was particularly true if the anger was expressed toward a parent. Schools, too, are resistant to emotional expressions. The net result is that most people do not learn appropriate ways to express emotions, or for that matter, to be with someone else who is expressing emotion in their presence. If the discrepancy is great between the way one perceives one's self as handling emotion producing situations and the way one sees a really adequate person handling it, the person will tend to bottle it up and seem passive. This same person, in a sufficiently provocative situation, may be unable to control it. The feeling will come out explosively, possibly even doing violence to others. On the other hand, a person whose perceived self is only slightly discrepant from his concept of adequacy, will easily express his warmth, his annoyance, his grief. He is more capable of showing feeling at any level of intensity, but he is less likely to show extremes of feeling, anger for example, because expressing annoyance, as it is experienced, keeps it from building to unmanageable proportions.

Autonomy

The fourth center around which behavior is organized is called autonomy. It is the feeling that one can make choices which will have a significant effect on one's life. We all start out life as highly dependent upon adults. As we develop some feelings of worth, ability to cope and to express, we learn that, in every situation, there are alternative courses of behavior. Some of these alternatives are more satisfying than others. If we were truly independent, we would be guided by our internal standards of evaluation and become more autonomous. However, our experience with authority, and the way in which others have manipulated the rewards and satisfactions, tends to oppose this development. There are many situations in which people feel, "I have no choice." This is not literally true, of course. One always has a choice, but choices have consequences. It may well be that, because of the consequences, even a highly autonomous person would make the same choice as others in a given situation. The feelings would be quite different, though. A person who says, "I have no choice," feels a lack of control over aspects of his life. The autonomous person feels that he always has a choice, and that the course of action he takes is his choice. He can consider alternatives with a real feeling that the issue is not foreclosed. Thus, the person who perceives himself very differently from his picture of adequacy may complain, but is docile with authority, feels controlled and helpless to change many aspects of his life. Where the discrepancy is small, the person appears to be independent and optimistic that, by his own efforts, he can better his situation.

These four organizing centers in the self concept are interrelated in many ways. The feelings of personal worth seem to be a kind of cornerstone of personality. It provides the pleasure and optimism which makes growing up exciting and worthwhile. One has the courage to try to cope with this complex world. But, trying to cope also brings frustration, or sometimes joy, so there is need to express. Our culture's prohibition of outwardly expressed strong feelings leads them to be turned inward or dammed up. When anger is turned inward it tends to lead to feelings of depression because anger at the self gives rise to punitive impulses toward the self for not being more capable. Despite this, if the individual is encouraged and loved, he finds that he can cope with the world, and this, supported by his feelings of worth, enables him to assert some autonomy. A worthwhile person with coping skills can make choices which he would otherwise fear or with which he would feel unable to cope.

As an individual grows, his perception of himself changes to include his accomplishments. His concept of adequacy also grows to present a more complex picture of what the individual could be like. If these two areas of self grow in such a way that the discrepancy between them is never too great, the person will appear appropriately mature for his age. As an adult, a person who is mature, according to this description, would see himself as a worthy person with very little need to defend this perception. He would move into various situations with the confidence that he would probably be able to cope effectively. His expression of feelings would be quite open and accurate, and he would feel that he could really make choices which would have some influence over his own destiny. Such a person would probably be similar to the description that Maslow (1970) has given of a self-actualizing person.

Motivation and Behavior

This is the heart of the model and will, I think, provide a basis for looking at the components of the learning environment which have an important influence on affect and affective development in students. However, I would like to sketch in the rest of the model very briefly, as I think it will clarify some additional things about the learning environment.

Motivation to maintain the self and, when the opportunity presents itself, to reduce the discrepancies between the perceived self in the world and the concept of adequacy, are present at all times. The activity of an individual at some particular moment brings into figure the motivations inherent in whichever is the appropriate organizing center in the self. While actually working on learning some skill, the activity is energized by the motivation to cope more adequately. Interpersonal relations, between boys and girls for example, in addition to the biological drive, would be energized by motivation to become more worthy. The painful struggling to find one's feelings, and talk about them, in an encounter group draws upon the motivation to express. Outbursts against authority, as unskillful and unlikely to succeed as they may be, are still motivated by the drive to become more mature through becoming more autonomous.

Over time, an individual has learned many ways of maintaining and enhancing the self. These are his strategies for becoming, which may include an orientation toward dependency, that is, a tendency to rely on authority. Another may turn from authority and try to find his own ways of becoming more adequate, a kind of problem solving orientation. In addition, people adopt short and long range goals which guide their choice of activities. One might decide to go out for the basketball team or select some vocational goal, such as becoming a teacher. Such goals are usually achieved by known patterns of behavior, such as practicing basketball skills, or attending a college which has a teacher preparation program. These strategies or goals have been incorporated in the self because they appear to offer routes toward greater adequacy.

All of the behavior I have discussed occurs in a context which is called, on the diagram, the action space. For our purposes, this would consist of the classroom and school environment. Within the action space, there are many choices an individual may make. Perhaps any behavior is theoretically possible. However, by various means, a particular action space, such as a classroom, delimits these choices. There are rules of conduct within a classroom; there are approved activities which the school believes will enable a child to achieve certain educational goals; and, there are selected resources provided. Thus, a math classroom is not normally the place within which a boy can practice his basketball skills or work too actively on developing his relations with his girlfriend. Not only does the classroom limit and direct choices, the self concept of the child and the strategies he knows, places limits upon the choices he can make. A child who sees himself as unable to express himself well may not feel he has the choice of volunteering to speak in class. If he sees himself as unable to cope mathematically, he may not feel that he can choose to do the assigned work. Perhaps, in his perception, being seen as lazy or defiant is preferable to being a known failure.

Despite the fact that a classroom is usually structured to encourage certain kinds of behavior and discourage others, a child may work on other goals. Much of the talking or whispering between children is concerned with interpersonal relationships and is energized by the motivation to feel more worthy through caring friendships. It should also be clear that various ways of structuring the action space can facilitate or inhibit the opportunity for a child to work on the adequacies about which he is concerned. More will be said about this later.

Once the individual has assessed the action space, he acts in some way calculated to either maintain the perceived self or enable him to feel more adequate. This model of the organization of the self has been inferred from observation and interpretation, over time, of the overt acts of individuals. What we observe is the content of the act. But, because we believe all behavior is caused, we try to understand the intent of the act as one part of the causal chain. The intent is inferred from examining the context, or action space, and giving attention to all aspects of the content of the act. It has both cognitive and affective content. We can observe the tone of voice, body tension, gestures, changes in skin coloring such as blushing, facial expressions, and the words which are used. When all of this has been considered, we ask ourselves, "Can we understand what the individual was trying to accomplish?" The answer we come up with will often determine the response we make, and this response then becomes the consequence, or one of the consequences, perceived by the child as the result of his act. The model which has been presented here has been inferred from behavior and should serve as an aid in determing the kinds of consequences which will relate to the child's motivation and, therefore, be most helpful to his growth. What we are concerned with in trying to improve our measurement of affect, is the kind of things which are important to sample from the action space, the individual's acts, and the consequences provided by the school.

Parenthetically, it is worth noting that this last part of the model, represented by the squares in the diagram, is the part of learning which is of concern to B. F. Skinner. Those behaviors which are wanted are reinforced; that is, because pleasant consequences occur, those acts, which we value, are performed by the individual. The rest of the model, which Skinner would reject, is concerned with understanding and predicting which act the individual will perform.

Aware–ing

Before we continue with the question of how and what to measure, there is one more part of the model which should be described. The consequences which the individual experiences after he acts are interpreted by him in terms of their meaning for the self. Because this interpretation is a somewhat complex process for which there seems to be no existing word, I have called it "aware-ing." It is a process in which the consequence of the individual's action is evaluated in terms of its consistency with the self concept. If it is consistent, then it can be accepted and evaluated a second time in relationship to adequacy. Does the consequence mean I have become more adequate? If the consequence is inconsistent with the concept of self, it is rationalized or denied. For example, if a teacher tells a student, who sees himself as average, that his work was satisfactory, he can accept this as consistent with his perceived self. Since the work is designed to make him more adequate in coping with some aspect of life, he feels that he is becoming more adequate. However, if, with the same student, the teacher gives him strong praise, this does not fit his conception of himself as average, and, while he will feel good about it, he will tend to rationalize it away: "The teacher was just in a good mood," or "I was lucky." We all know how difficult it is for many people to hear a compliment. Because the student has explained away the compliment, it tends not to make him feel more adequate. The same thing is true when the teacher gives the student negative feedback; he will probably rationalize it and, therefore, discount it. If the feedback over time is consistently different from his earlier belief, there will be a gradual change in perceived self and, as a result, in his feelings of adequacy. The kind of consequences which capable teachers provide in the classroom are typically of this kind, that is, a gradual building of feelings of confidence in coping or of personal worth, and so forth. It tends to be gradual because the aware-ing process is like a channel which will allow in only those things which are consonant with the perceived self. Other consequences arouse negative feelings or strong emotions or anxiety and lead to behavior which defends the self from such input. Rogers (1959) speaks of a more rapid kind of change. He sees the organization within a person as a kind of gestalt, where it is possible, by changing a minor item, to alter the whole pattern of organization. This sometimes happens when a student goes to college. He has been an excellent student in high school but has his confidence quite suddenly shattered in college when he discovers that his typical efforts now only rate a mediocre grade. In competition with the best from other high schools, the feedback quickly makes him feel far less adequate. This kind of change seems to be rare in the schools. It sometimes happens when an average student suddenly develops a deep interest and, with a stronger motivation, begins to turn in outstanding work. The change usually is puzzling but it is probably because the student suddenly sees the work as really enabling him to cope better with something important to him.

Learning

The diagram in Figure 2 can be divided into two parts: that illustrated with circles and that presented as squares. Change can take place in any part of the system, and the new configuration will result in different behavior. If a change takes place in one of the squares, that is, if an individual perceives his action space differently, or learns new skills or knowledge that change his actions, or if he interprets consequences differently, the change is one which I call instrumental learning. The self content is not changed necessarily. However, when an individual takes on new identifications in his concept of adequacy, there is now a change in the discrepancies between the perceived self and the concept of adequacy. Thus, motivation changes and, quite possibly, changes in the strategies for becoming or in the person's goals also take place. I call this intrinsic learning. They are clearly interrelated, but they also differ in important ways. If the teacher pays attention only to fostering actions which represent skills and knowledge, the process is essentially training. Available choices in the classroom and the teacher's actions which provide the student's consequences are directed at cognitive outcomes or motor outcomes. The teaching of typing represents a simplified example. The classroom is designed for this one function and the procedures are straight forward. Students are instructed to put their fingers on the keys and to perform various actions. As long as the student does as he is told, the teacher has little concern with why the student wants to type, or what he will use it for when he gets through. If we pay attention to only the self concept part represented by circles, then our concern is with therapy. When the discrepancy between the perceived self and the concept of adequacy is decreased, the person will function more effectively, that is, have better interpersonal relations, cope more confidently, express himself more openly, and take more responsibility for his own life. It seems clear to me that sound education is concerned with both parts of the model. We want students to build the necessary skills and knowledge but we also would like to insure that they use them effectively.

AFFECT AND THE LEARNING ENVIRONMENT

This model suggests that the learning environment has an impact on the cognitive-affective elements which make up the self concept. These are the factors which have significant effects upon an individual's feelings about himself. I would like to identify some of these briefly and discuss ways in which they can be measured. In terms of the Figure 2 diagram, I am concerned with the squares: the action space, the student's acts or behaviors, and the consequences provided by the school.

In a general sense, the action space perceived by a child includes all of the matters which influence him. It would include his parent's attitudes, possibly the public's reaction to the schools, the general climate created by administrative policies and rules, the size of the school, the number of children in the classroom, the design of the curriculum, and so forth. Even the aesthetic aspects of the grounds and interior decor have their impact. More directly, the expectations of the teacher and ways he or she relates to children and the relationships among the children are of critical importance. The most important of these influences are, perhaps parental attitudes, the perceived relevance of school and the material to be learned, the child's perception of the teacher, and the child's relationship with other children. Table 1, while far from complete, is an attempt to list some of these factors under the organizing centers of the self concept which they would influence. The quality of these factors in a classroom could be sampled either by direct observation or by an inventory completed by the students. Both techniques will be described below.

In a small volume entitled, *A System for Assessing Affectivity*, Bob Bills (1975) has developed instruments for sampling some of these. His Parent Inventory is a 35-item multiple choice instrument which samples parental attitudes and opinions. The first four items gather information about the grade of their child in school, their citizenship, and their levels of education. These items are not scored. The rest of the inventory concerns feelings about the quality of the school, the teachers and the way they handle the children, and the kinds of things which the parents think the school should be helping the children to achieve. The inventory yields an attitude score from negative to positive feelings about school. Bills' studies have shown that it is a reliable instrument which is able to differentiate between grade levels and among schools.

A second of Bills' instruments is a fairly global sampling of many aspects of the school, named, Feelings About School. There are two forms of this inventory with 50 items in each. The areas sampled have to do with the relevance of school work, the fairness of the school, how students think the teacher feels about the children, the children's feelings about the teacher, how interesting schoolwork is, how they feel about grading, and further questions of this kind. It, too, is simple to score, has adequate reliability, and shows differences by sex, grade level, and among schools. It has been used from 6th through 12th grades.

While some attitudes toward the teacher are sampled in the Feelings About School instrument, Bills has also developed a Relationship Inventory. This has 72 items and enables students to describe the qualities of the relationship they have with their teachers. The statements read, for example, "He (she) respects me," "He (she) tries to understand exactly how I see things," "He (she) is acting a part with me." It yields four scores on interpersonal relationship variables. These are based on Bills'

TABLE I

Elements of school organization, of the curriculum or of teacher behavior which are likely to have impact on a person's feelings of personal worth, confidence that he can cope, feelings that he can express openly and accurately, and that he can make significant choices.

Worth

Warm-cold or impersonal climate
Respectful – sarcastic
Strengthening – belittling
Non-punitive – punitive
Non-judgmental – judgmental
Valuing individual differences
Empathetic

Coping

Challenging – dull
Encouraging – discouraging
Flexible pacing – lockstep
Appropriate ratio of concrete/abstract
Flexible sequencing
Differences of opinion valued

Expressing

Acceptance of feeling expressions
Facilitation of feeling expressions
Modeling/Teacher expresses openly

Autonomy

Genuine choice opportunities
Clear, non-dogmatic rules
Risk-taking encouraged

theory of teacher variables which influence personal growth of students. Adapted from Rogers (1961), the variables are: empathic understanding of students, positive regard for students as important people, positive regard offered without conditions, and congruency in their relationships with students. To quote from Bills:

> The term empathy is used here to describe a relationship in which the student believes his teacher understands what it is like to be a student – the teacher can 'walk around in the shoes of the student.' Positiveness of regard merely means that, from the point of view of the student, the teacher acts as if he believes the student is an important person. If a teacher is unconditional in his regard for a student, it seems to the student that the teacher is not placing a 'price' on his positive regard, saying in effect, 'I will regard you highly if you are the kind of person I value and will not regard you highly if you are different.' And by congruence it is meant that the student perceives his teacher to be thinking and feeling in a manner consistent with how he is behaving – the teacher does not give the appearance of thinking and feeling one way and talking or behaving as if he is thinking and feeling in another way. (1975, p. 42)

The students respond to the items of the inventory by using a 6-point Likert-type scale. Bills has computed the reliability of the scales. The intercorrelations among them runs from .74 between empathic understanding and level of regard, to .34 between unconditionality of regard and level of regard. The factor of unconditionality of regard seems to be the most independent. Bills suggests that these factors have considerable separateness but are often present in the same person to the same degree. I would expect that unconditionality of regard is the one which most people are least prepared to give. We tend, as I said earlier, to get coping behavior entangled with the rewards of approval and love. In fact in Bills data, the lowest scores in one study, with an N=6929, are on this factor. The scales showed adequate split-half reliability and individual item correlations with total scale. The four scales give significantly different mean scores and show differences by sex, grade level, and school. The data which Bills has gathered also show that students rate their teachers as having higher regard and congruence than empathic understanding or unconditionality of regard. Male students scored their teachers lower on all four variables than did female students. Unconditionality of regard and congruence scores stay about the same from lower to higher grade levels, but scores on level of regard and empathic understanding decrease. This is interpreted to mean that as students proceed to the higher grades, they find that teachers regard them less highly and are less empathic.

On these three inventories developed by Bills, all of the scores indicate the way in which the school setting and teachers are perceived on scales which run from negative to positive. He has tested samples totaling from 2,000 students or parents up to more than 6,000, from grades six through college, and from six to ten different schools, depending on the instrument. There are neither norms nor clear means for interpreting results. It also seems possible to fake such inventories in order to create a good impression. Bills has collected his data with promised anonymity and there is certainly no need to know individual identities in order to assess the learning environment.

The research by Fox et al. (1964), referred to earlier, showed that relationships with classmates are also important. Two variables are of particular interest: how well a child is liked and how broadly diffused friendships are over the class. Better liked children made better use of their abilities, and the more children in a classroom who had friends, the more children who made fuller use of their abilities. The techniques for sociometric measurement are well developed and can be carried out quite simply in the classroom. Each class member indicates his choices of other children he or she would like to be with. The question could be phrased in terms of a number of settings such as choice of seatmates, classmate to invite home, classmate as a fellow committee member, and so forth. Usually a child is asked to make three choices and it is quite common and useful to ask them to indicate three whom they would rather not have with them. The data can be displayed as a diagram with arrows connecting children to show the pattern of choices and rejections.

It can also be organized as a sociograph (Clark, 1952), which shows status among peers by indicating those most chosen and those chosen by popular children. It also reveals the lines of communication in the class. Cleavage lines appear when none of the children in a higher status group have chosen any of the children in a lower status group. Not only does this mean that there are different status rankings, but that one group literally does not talk to the other group. In terms of Fox's research, the more cleavage lines, or the more rejected children there are in a classroom, the fewer children who will be fully utilizing their abilities.

With these examples of how the action space can be measured, I would like to move to the next area in my diagram, the student's actions in the action space. I shall make a few general comments and combine further consideration of the student's action with the next area – the consequences a child experiences in the classroom when he acts.

Attitudes toward subject matter and the values which students hold are relevant to learning. Teachers often have an objective to develop positive attitudes and clearly, attitudes, can facilitate or block learning. Inventories for measuring them are well developed and self devised scales, using Remmers or Likert type scales, are easy to use. The means of student scores on these gives an indication of classroom impact and positive attitudes tend to correlate with retention and possibly with future use of skills or knowledge. This probably needs no further discussion here.

The interdependence of a child's behavior and the consequences which follow is a natural unit of measurement and it reveals what is being reinforced. However, it has been difficult to gather this data. It seems increasingly evident that to split behavior and consequences apart simplifies the data to the point of meaninglessness. It could well be the reason that so many earlier studies in education have come up with the result "no significant differences." Soar (1972) states it this way, "Growing evidence for the existence of complex interactions between classroom behavior, the nature of the learning task, and characteristics of pupils, points strongly to the need for more subtle, more complex, and more extensive research on classroom learning." (p. 23). One technique which seems to meet some of these needs is observation. It has some disadvantages in that it requires training of observers, takes the time of a skilled observer, and may, at least initially, make teachers and children self conscious. It does have definite advantages, however. It gives data based directly upon observed behavior. It keeps the connection intact between the learning task, the teacher behavior, and the student behavior. Finally, it makes possible the sampling of a more complete range of factors which influence pupil growth.

The development of satisfactory observation instruments is still continuing, but there is one instrument which has been widely used and has probably been subjected to more study and development than any other currently available. This is Ned Flanders' Classroom Interaction Analysis (1970). In a book published in 1970, *Analyzing Teacher Behavior*, Flanders has discussed classroom behavior techniques for recording, analyzing and interpreting the data, and a number of uses which can be made of the technique. Since this resource is available, I will limit the discussion here to the process of data collection, the kind of data that it produces, and a few of the major findings from its use.

The term "interaction analysis" refers to any technique which attempts to gather data on the chain of events in a classroom. A trained observer sits in the classroom and records the flow of events according to predetermined categories. He learns code symbols for each category, and as he observes the events in the classroom, he writes down the code symbols which classify the event. With this long list of categorizations, it is possible to analyse the frequency of events in each category and to see how each event is part of a longer chain. Inferences about the chain then can be made. It is important to make entries at a consistent rate of 20 to 25 tallies a minute in order to know what proportion of the time is spent in each activity. Flanders' approach is a ten-category system (Table 2). Seven categories are used when the teacher is talking, two when students are talking, and one to indicate silence or confused activity.

Of the seven categories of teacher talk, the first three are responses to students: "accepts feelings," "praises or encourages," and "accepts or uses ideas of pupils." A fourth category is "asks question." In early research, these four categories were referred to as indirect teacher influence. The next three categories, "lecturing," "giving directions," and "criticizing

Table 2. Summary of Categories for the Flanders System of Interaction Analysis

TEACHER TALK	INDIRECT INFLUENCE	1. Accepts Feeling: accepts and clarifies the feeling tone of the students in a non-threatening manner. Feelings may be positive or negative. Predicting or recalling feelings are included.
		2. Praises or Encourages: praises or encourages student action or behavior. Jokes that release tension, not at the expense of another individual, nodding head or saying "um-hm" or "go on" are included.
		3. Accepts or Uses Ideas of Student: clarifying, building, or developing ideas or suggestions by a student. As teacher brings more ideas into play, shift to category five.
		4. Asks Question: asking a question about content or procedure with the intent that a student answer.
	DIRECT INFLUENCE	5. Lecturing: giving facts or opinions about content or procedure; expressing his own ideas, asking rhetorical questions.
		6. Giving Direction: directions, commands, or orders to which a student is expected to comply.
		7. Criticizing or Justifying Authority: statements intended to change student behavior from non-acceptable to acceptable pattern; bawling someone out; stating what he is doing; extreme self-reference.
STUDENT TALK		8. Student Talk Response: talk by students in response to teacher. Teacher initiates the contact or solicits statement.
		9. Student Talk Initiation: talk by students which they initiate. If "calling on" student is only to indicate who may talk next, observer must decide whether student wanted to talk. If he did, use this category.
		10. Silence or Confusion: pauses, short periods of silence and periods of confusion in which communication cannot be understood by the observer.

From: Flanders, N. A. *Teacher influence, pupil attitudes, and achievement.* (Cooperative Research Monograph No. 12, OE 25040) Washington: U.S. Government Printing Office, 1965.

or justifying authority," were called direct teacher influence. Flanders' research showed that in classes in which students scored high on liking the teacher, feeling motivated, believing there were fair rewards and punishments, lack of anxiety, and independence, the teachers used more indirect influence; whereas, teachers of the classes in which students scored low on these factors used more direct influence. The differences between teachers classified as using more indirect influence versus more direct influence are small. Flanders comments, "One gets the impression that a small amount of indirect influence lubricates the classroom gears of subject matter learning, and even though the total amount of indirectness is small, its presence or absence is significantly related to the positive or negative attitudes of pupils, respectively" (1970, p. 392).

One other instrument which produces data about some of the interactions in the classroom is a further one developed by Bills (1975) entitled, Locus of Responsibility Scale. It is a 27 item multiple choice inventory designed to tap the students' perceptions of the nature of classroom interaction. As Bills describes it, it attempts to answer the question: "Who is making decisions in the classroom?" The answer comes as scores for each of three categories: teacher-centered, interactive, and student-centered. Bills suggests that scores that are balanced between the three categories, with the highest score on interactive processes for making decisions, seem optimal. This instrument differentiates among grade levels and schools and between sexes. Although both sexes see teachers about the same way on teacher-centeredness, females report higher student-centered relationships. There are many possible patterns of interaction, some of which might correlate with affective variables such as attitudes toward school. Bills does not report any such studies.

Measuring Affect in Individuals

The self concept model which I have presented becomes even more useful when one looks at affective processes in individuals. Specifically, the measurement of individual affect concerns the kind of cognitive-affective "acts" of individuals which enable us to infer feelings about adequacy. How can we measure the cognitive-affective dispositions associated with the organizing centers of worth, coping, expressing, and autonomy or choice making?

In order to answer this, I have followed a strategy proposed by Jim Popham (Payne, 1974) for specifying affective objectives. It consists of five steps:

1) Make a general statement of the affective objectives.
2) Imagine a student who personifies these objectives and try to describe the behavior which he would exhibit.
3) Imagine a student who does not possess these qualities, or possesses the opposite feelings, and try to describe the behavior he would exhibit.
4) Select situations in which the two imagined people would respond differently. This could be a naturally occurring situation or one that is contrived.
5) As a final step, one should select those situations which most effectively, efficiently, and practically define the intended outcomes.

A general statement of the affective objectives might be as follows: As students progress through school, they will maintain or increase feelings of personal worth, of being able to cope, of being able to express openly and accurately, and of being to some degree autonomous, that is, able to make their own choices.

The statement uses the words, "maintain or increase." The assumption is that some children have relatively healthy self concepts; that is, there are small discrepancies between the perceived self and the concept of adequacy, and the classroom should help them maintain this. Others, with larger discrepancies between perceived self and the concept of adequacy, will have classroom experiences which will enable them to increasingly feel that their perceived selves are becoming more adequate.

The attempt to follow steps two and three in the Popham strategy are presented in Table 3. The behavioral characteristics listed for students who have feelings of adequacy and those who feel much less adequate can serve as a guide for determining the kind of data one needs to gather, or for the selection of an instrument which will measure dimensions of the self concept.

Once again, the observation could be used to look at the behavior of individuals. Though it may be an effective method, it certainly is not the most efficient or practical. For research purposes, observation could serve as an independent measure for

TABLE 3

The cognitive/affective variables of worth, coping, expressing, and autonomy for two hypothetical students.

POSSESSOR	NON-POSSESSOR
Worth	*Worth*
Positive self reference (self accepting)	Seeks reassurance/denies own importance
Comfortable with people (poised)	Nervous around people
Is friendly (has friends)	Few friends/quarrelsome with peers
Can admit imperfections (take criticism)	Boasts
Independent	Dependent on authority or others
Flexible	Rigid
Undefensive/open	Easily defensive/denies actions
Trusting	Suspicious/wary
Warm (loving)	Cold
Spontaneous	Calculating/mean
Can relax (not tense)	Often tense/rarely relaxed
Happy	Unhappy
Can be quiet	Attention getting

Coping	*Coping*
Takes responsibility/good leader	Poor leader/irresponsible
Sees school work as challenging	Dislikes work
Curious/participates	Dislikes novelty/reluctant to participate
Problem solving approach	Flounders
Likes school	Dislikes school
Competent in school work	Does poor work
Sociometrically chosen for tasks	Not chosen for tasks

Expressing	*Expressing*
Talks easily in class/outspoken	Does not volunteer to participate
Shares feelings (even negative)	Hides feelings/pokerface/fake smile
Takes positions on issues	Not clear what he believes
Open about feelings/self revealing	Denies feelings
Direct/blunt	Super tactful/unclear meanings
Constructive confrontations	Destructive/blows up/blasts
Less superficial talk	Chatters

Autonomy/Can Make Choices	*Autonomy*
Can disagree openly	Never disagrees or really agrees
Decisive	Wish-washy
Can take reasonable risks	Plays it safe
Thinks through issues & decisions	Impulsive/flip-flops
Resistance to authority/reality oriented/tests for self	Authority oriented
Internal locus of evaluation	Dependent/checks out others

validating self report type measures. It is also a useful tool in the case of an individual child who seems to be having difficulty in school. In this case, rather than an interaction analysis type observation, it would probably be most helpful to employ an anecdotal record as described by Prescott (1957).

When we turn to other measuring instruments, it is clear that there are problems which do not appear when measuring cognitive variables. The first problem is that, with the exception of some attitude inventories, very little affective measurement has been done in the schools. There may even be resistance to taking the time for such measures since most of the existing ones, developed independent of an explanatory theory, have not supplied data which had clear usefulness. They may be useful directly to counseling situations, but for the most part, they do not seem to get very close to the total dynamics affecting motivation and learning. It seems to me that those of us concerned with measurement and research in education must help teachers understand the importance of affective factors.

A second kind of problem, related to the first, is that some of the better known psychologists, Skinner (1974) for example, deny the importance of internal processes. These, too, are seen as the resultant of reinforcement processes. They argue that, if one provides the appropriate consequences skillfully enough, any behavior one desires can be developed. I do not believe that behaviorists have demonstrated this with the more complex human processes. At the same time, I think there is evidence that internal variables correlate with behavior and predicted outcomes. I have already reviewed some of this research. In the end, it is essentially an empirical question to be settled by further research.

A different kind of problem arises because, with the exception of observation, affective measuring instruments are, for the most part, self report statements. It is clear that these are subject to distortion. In a study by Getzels and Walsh (1958), it was assumed that any question serves as a stimulus for an individual to set up a personal hypothesis as to how he would answer the question. He then checks his personal hypothesis with the situation. Following that, a verbal response is made which is either the personal hypothesis or a distortion which fits the situation. With this mini-theory, they put forth several hypotheses for research. The first one was that questions dealing with socially conflicted objects of inquiry would elicit a large discrepancy between the personal hypothesis and the expressed reaction. The second hypothesis stated that questions dealing with socially neutral objects would elicit small discrepancies between the two. They tested the hypotheses by giving persons two questionnaires, one made up of direct questions, and one, a sentence completion questionnaire, which was scored as a projective technique. Both of their hypotheses were confirmed.

The implications of this study are essentially that, in any attempt to measure affective variables, we are asking students to make public statements subject to a social desirability effect. In the cases where such a public statement might invoke defensiveness, techniques should be employed which minimize this. In Getzels' study, the use of a simple projective technique was able to circumvent much of the distortion. It is also possible to use a paired comparison technique in which statements are selected which have a low correlation with scales of social desirability, such as those developed by Edwards (1957).

There are problems connected with establishing validity for affective instruments. Reliability and content validity seem to cause few problems. Concurrent validity, the correlation between the instrument and independent measures of the same trait, presents the problem of finding really good independent indicators. Correlations with other tests, purportedly measuring the same construct, are usually positive, often significant, but low Construct validation, the degree to which an instrument accurately measures the construct, presents problems which I think are mainly due to the difficulty of defining the construct. I think that improvement in this area and in the area of predictive validity will come as we develop definitions more tied to behavior. I have tried to do this in Table 3. Certainly it is an area which calls for much work and ingenuity.

Warren Findley, in a personal communication, called my attention to a paper submitted for presentation at AERA by Walker and McGranahan (1976). The paper concerns the development of an affective measure which samples four kinds of attitudes: feelings toward school, social structure and climate, attitudes toward reading and math, and self concept in school. An especially interesting part of the study was the validation technique used. A group of students was asked to play-act that they enjoyed school and felt competent in their school work. A second group was asked to play-act the opposite. To quote from Dr. Walker's letter: "The logic of the study was this: if students who are pretending to have a specific attitude give a certain set of responses, then those responses are a valid indicator of the attitude when given by students who are not role-playing." In addition to the imaginative technique, there was an interesting finding. Students role-playing positive attitudes achieved very high scores on the attitude measures, while those role-playing negative attitudes had scores close to zero. This clearly indicates that what the test constructor thought were items which would measure positiveness and negativeness agreed to a very high degree with students' perceptions. This would seem highly validating. However, "real" students do not score at these extremes, most likely because their feelings are mixed or qualified in some way. Since constructs are defined in "pure" terms (a person has it or does not have it), the results of validation studies do not come out as clearly as we would like.

The solution to validity problems still lies ahead, but we should be aware that we are working with constructs describing very complex processes and functions. "There is no royal road to 'self;' we are forced to approach along the only paths open to us, through the tortuous byways of analysis, inference, hypothesis, and reconstruction" (Mason, 1976). The prize is increased ability to predict and to formulate effective teaching methods.

Despite the problems which we have discussed, there are literally hundreds of affective measures. Some are quite well known such as the Rorschach or Thematic Apperception Technique. A large number of newer ones are briefly described in *Improving Educational Assessment and An Inventory of Measures of Affective Behavior* (Beatty, 1969) under headings such as "Attitude Scales," "Creativity," "Interaction," "Motivation," "Personality," and "Self Concept." The measures are in various stages of development and offer a variety of approaches to the measurement problem.

Before talking about a specific instrument which has been developed just for the purpose of assessing individual affectivity in the schools, I would like to give brief descriptions of several approaches to affective assessment. One technique which is not too well known is the Semantic Differential, developed by

Osgood, Suci, and Tannenbaum (1957). They assume that adjectives are the most important words for conveying the shades of meaning which concepts have for us. They experimented with a number of bi-polar adjective scales for rating concepts. Of the many different dimensions studied, three were found to account for most of the variance. They were evaluation, for example, good-bad; potency, that is, strong-weak; and activity, adjectives such as fast-slow. Actually, the major part of the variance is accounted for on the evaluation dimension alone. This is probably true because it is the main element of affective reactions. If a concept such as school was studied, students might be asked to rate it on a number of 7-point scales concerned with evaluation such as, good-bad, beautiful-ugly, fair-unfair, happy-sad, and nice-awful. If ten scales were used, one would get scores on a 7-point scale ranging from ten to 70. These scores could then be compared to determine positiveness or negativeness toward various aspects of school, differences in attitudes between the beginning and end of a school year, across grade levels, and so forth. The concepts could, of course, be chosen to sample aspects of worth, coping, expressing, and autonomy.

Another measuring tool, appropriate to affective variables, is the Q-sort developed by Stephenson (1953). The technique was used by Rogers (1954) to assess personality changes associated with therapy. The Q-sort is constructed by selecting a number of statements about personality or personal behavior. It is possible to use a Q-sort developed by someone else, but in order to include those particular behaviors which are relevant to the study, the items are usually developed by the person doing the assessment. These statements are printed on separate cards, and the person being tested is asked to sort them into piles ranging from "most like me" to "least like me." In order to insure uniform distribution of the traits, a "forced-normal" distribution is used. The subject is asked to place a specific number of cards in each pile. For example, if the sort had 54 items in it, one would use nine piles and place only three items in each of the extreme piles, five in the two next most extreme, six in the next, eight in the next, and ten in the middle pile. The data derived from this procedure consists of an individual's beliefs about his characteristics. It can also be used in a number of other ways such as sorting the items the way one thinks the teacher sees one, or in terms of one's concept of adequacy. An individual can also sort according to the way he thinks most descriptive of another person. Each of these sorts represents the sorter's perception, and there are no norms which can be applied. Comparisons can be made by computing correlations between two separate sortings of the items. It is also possible to discover which items are used most often within a group of students to describe themselves or others. Stephenson describes even more elaborate ways of constructing and using the method. The scoring is somewhat time consuming, but the instrument has great flexibility and could certainly be used as an independent criterion while studying the validity of another measuring device.

Projective techniques are particularly appropriate when there is concern that students might want to distort their answers to create a particular impression. Some projectives are complex and take long training to score. There are simpler approaches, however, such as the sentence completion technique mentioned in Getzels and Walsh's (1958) research. A person is presented with part of a sentence such as, "People like me . . .," and he is asked to finish the sentence in any way he wishes. It is possible to look at the content of the ending provided by the person for cognitive reactions, but as an affective instrument, it is scored in terms of whether or not the completion shows positive, negative or indeterminate affect. In the case of the stem, "People like me . . .," if the student finished it with the phrase, "most of the time," it would be scored as positive affect. It is possible to get high agreement among raters and this approach has great flexibility. It has the usual advantage of a projective technique, that of being able to tap feelings which might be hidden if asked for directly. By measuring variables both directly and with this projective technique, the investigator could use the discrepancy between responses as an indication of defensiveness. High defensiveness is suggestive of doubts about personal worth.

The Index of Adjustment and Values, developed by Bills et al. (1951), is a self perception measure which has probably been more thoroughly studied than any other single instrument for assessing self concept. It began in 1951 as a measure of adult adjustment, and through continual work, now offers forms to measure adults, high school, junior high school, and elementary school students from the third grade on up. Each of these forms has been tested for reliability and, to some degree, for validity. The adult form has also been used extensively in research. The instrument, the validation studies on it, and some of the research results using it, are reported by Bills in his book, *A System for Assessing Affectivity* (1975). I would like to present some of the highlights from that report.

The instrument consists of two answer sheets, each containing a list of 49 adjectives, and two instruction sheets. The Index of Adjustment and Values (IAV) asks an individual to rate himself on a five point scale with regard to how much of the time the adjective characterizes him. He is then asked to rate, again on a five point scale, how he feels about being this way, and finally, to rate how he would like to be with regard to the same trait. When he completes the ratings for himself he is then asked to make the same assessments as he believes his peers would rate themselves. The two answer sheets are identical. The instruction sheets differ only in that the word "he" instead of "I" is used to indicate that one is rating others rather than self.

The instrument yields scores showing the difference between the perceived self and the perceived ideal self, a comparable score for perceptions of other people, and two scores which summarize the relationship between a person's self acceptance and his perception of other people's self acceptance. This is done by summing each of the columns and computing a discrepancy score between Column I ("I am ______") and Column III ("I would like to be ______").

The IAV for Self can be examined at several levels. Each rated trait gives specific information about how an individual sees himself, how he feels about it and how he would like to be. Such information is probably of use only in counseling. It is also possible to compare scores with normative data provided by Bills. This would probably be most useful in examining or comparing group means. For example, is this class more self accepting than that class? The third level would be concerned with the behavioral characteristics of people with high or low scores on the IAV. Bills provides some information on this. He suggests that self acceptance is not a linear function but one in which scores around or just above the population mean represent optimal adjustment.

Bills has developed another score which appears to be a real innovation. This is what he calls categorical scores. By using the Column II ("How you feel about yourself") sums from both the Self and Others ratings, a person can be assigned to one of four categories. If his Self Column II score is at or above the population mean, he is assigned a plus (+). If the rating score for Others is equal to or above the rating he gave to himself, he is assigned another plus, and is labeled a ++ person. If a person's Column II score for Self is below the mean for the population he is assigned a minus (–). If, at the same time, the ratings on Others is above his score on Self, he is given a second sign of plus. The four combinations then, make it possible to have categories of people labeled ++, –+, +–, and ––. Thus, ++ people are those who score themselves above the mean of the population on self acceptance and see others as equally or more self accepting. Plus-minus people see themselves as more accepting than the population but see others as less accepting of themselves. The –+ person is less self accepting but sees others as more self accepting and the –– people are neither self accepting nor do they see others as self accepting.

Some of the score interpretations which Bills believes are supported by his research are as follows. Increases in Column I and II on the Self measure are usually indicative of positive change. At the same time, a decrease in the discrepancy score between Columns I and III may be positive. Bills theory assumes that personally well-adjusted people like themselves and see themselves as little discrepant from their ideal selves. When it comes to category scores as a means of assessing change, two variables are taken into account together: perceived self acceptance and perceived self acceptance of others. Changes in categories are seen as indicative of change in self perception. To evaluate this, the four categories are interpreted from most desirable to least desirable in the order ++, –+, +–, and ––.

Research which supports these interpretations and data on reliability and validity are all reported in Bills' book (1975).

With the instruments which I have discussed for measuring affective variables in individuals, it should be possible to assess the four cognitive-affective organizing centers presented in my self concept theory. When measures of the learning environment are studied in relation to student scores on affective variables and their performance in school, it should be possible to deepen our understanding of the factors which enable children to learn in school and become effective people in their lives beyond school. Both sets of measures, those for the learning environment and those to assess students, have been drawn in relation to an overall theory of motivation and learning. This should enable us to develop an appropriate rationale for our measurement efforts, and the results of our measurement should enable us to correct and extend the theory. If such an endeavor lives up to its promise, it could well lead to important changes in the schools.

BIBLIOGRAPHY

Beatty, W. H. and Clark, R. A., A self concept theory of learning. In H. C. Lindgren (ed.), *Readings in Educational Psychology*. New York: John Wiley & Sons, 1968.

Beatty, W. H. Emotion: The missing link in education. In W. H. Beatty (ed.), *Improving educational assessment & an inventory of measures of affective behavior*. Washington: Association for Supervision and Curriculum Development, 1969.

Beatty, W. H. Validation and uses of the observation for determining the climate of feelings in a classroom. Unpublished doctoral dissertation, University of Chicago, 1952.

Bills, R. E. *A system for assessing affectivity*. University, Alabama: The University of Alabama Press, 1975.

Bills, R. E., Vance, E. L. & McLean, O. S. An index of adjustment and values. *Journal of Consulting Psychology*, 1951, *15*, 257-261.

Clark, R. A. Sociógraphic analysis of adolescent peer status. Unpublished doctoral dissertation, University of Texas 1951.

David, R. E. Teacher behavior: Its effect on student attitude change. Unpublished doctoral dissertation, University of California, 1976.

Edwards, A. L. *Techniques of attitude scale construction*. New York: Appleton-Century-Crofts, 1957.

Flanders, N. A. *Analyzing teacher behavior*. Reading, Mass.: Addison-Wesley Publishing Co., 1970.

Flanders, N. A. *Teacher influence, pupil attitudes, and achievement*. Coop. Research Monograph No. 12, Office of

Education 25040, Washington: U.S. Government Printing Office, 1965.

Fox, R. S., Lippitt, R. O., Schmuck, R. A. et al. *Pupil-teacher adjustment and mutual adaptation in creating the classroom learning environment*. Final Report, Coop. Research Project No. 1167. Office of Education, Washington: U.S. Government Printing Office, 1964.

Getzels, J. W. & Walsh, J. J. The method of paired direct and projective questionnaires in the study of attitude structure and socialization. *Psychological Monographs*, 1958, *72*, (1, Whole No. 454).

Gordon, I. J. Affect and cognition (A reciprocal relationship). *Educational Leadership*, 1970, *27*, 661-664.

Kelley, E. C. The place of affective learning. *Educational Leadership*, 1965, *22*, 455.

Krathwohl, D. R., Bloom, B. S. & Masia, B. B. *Taxonomy of educational objectives. Handbook II: Affective domain*. New York: David McKay Co., 1964.

Maslow, A. H. *Motivation and personality*, (Rev. Ed.). New York: Harper & Row, 1970.

Mason, W. A. Windows on other minds. *Science*, 1976, *194*, 930-931.

Osgood, C. E., Suci, G. J. & Tannenbaum, P. H. *The measurement of meaning*. Urbana, Ill.: University of Illinois Press, 1957.

Payne, D. A. *The assessment of learning: Cognitive and affective*. Lexington, Mass.: D.C. Heath & Co., 1974.

Piaget, J. & Inhelder, B. *The psychology of the child*. Translated by Helen Weaver. New York: Basic Books, Inc. 1969.

Prescott, D. A. *Emotion and the educative process*. Washington: American Council on Education, 1938.

Prescott, D. A. *The child in the educative process*. New York: McGraw-Hill Book Co., 1957.

Rogers, C. R. A theory of therapy, personality, and interpersonal relationships, as developed in the client-centered framework. In Sigmund Koch (ed.), *Psychology: A study of a science*, Vol. 3. New York: McGraw-Hill Book Co., 1959.

Rogers C. R. & Dymond, R. F. *Psychotherapy and personality change*. Chicago: University of Chicago Press, 1954.

Rogers, C. R. Significant learning: In therapy and in education. In C. R. Rogers, *On becoming a person*. Cambridge, Mass.: The Riverside Press, 1961.

Rogers, C. R. The significance of self-regarding attitudes and perceptions. In M. L. Reymert (ed.), *Feelings and emotions*. New York: McGraw-Hill Book Co., 1950.

Skinner, B. F. *About behaviorism*. New York: Vintage Books, 1974.

Soar, R. S. Achieving humaness: Supporting reseach. In N. K. Hamilton & J. G. Saylor (eds.), *Humanizing the secondary school*. Washington: Association for Supervision and Curriculum Development, 1969.

Soar, R. S. Teacher behavior related to pupil growth. *International Review of Education XVIII*, 1972, *4*, 23.

Stephenson, W. *The study of behavior: Q-technique and its methodology*. Chicago: University of Chicago Press, 1953.

Tyler, R. W. Outlook for the future. *Educational Evaluation: New Roles, New Means, 68th Yearbook of the National Society for the Study of Education*, Chicago: University of Chicago Press, 1969.

L

Walker, C. B. & McGranahan, P. Validation of a measure of students' attitudes (in preparation). Center for the Study of Evaluation, University of California at Los Angeles, California. 1976.

Withall, J. The development of a climate index. *Journal of Educational Research*, 1951, *45*, 93-99.

NSPER: 76

Co-Directors' Report

Dr. William J. Gephart

Dr. Robert B. Ingle

EVALUATION AND THE AFFECTIVE DOMAIN

William J. Gephart
Phi Delta Kappa
&
Robert B. Ingle
Univ. of Wisconsin-Milwaukee

There are two ways in which evaluation *and* the affective domain must be considered. The first is evaluation of affective aspects of programs, personnel, or products. The second is the use of affect *in* evaluation. Understanding the first encompasses (1) the general nature of the evaluation process, (2) what it consists of, and (3) how it works. Understanding the latter requires (1) knowledge of the affective domain itself and (2) the acceptance of an epistemology, a way of knowing, that is not held in high respect in our culture. This paper presents thoughts about these five points that were examined and re-examined at the NSPER: 76 sessions.

THE GENERAL NATURE OF THE EVALUATION PROCESS

Evaluation is a systematic problem-solving process. This is a basic point on which all of the evaluation theorists agree. As such, evaluation is a derivative of the scientific method akin to the research, the development and other systematic problem-solving processes. Like research and development, the evaluation process involves a problem, the design of a systematic plan of action to resolve that problem, the reliance on data in solving the problem, and the evaluation of a product. Evaluation differs from research and development and these differences became apparent when one specifies the general characteristics of the problem, the action plan, and the product for each of the processes.

To make clear the differences between an evaluation problem and a research or development problem, it is necessary to define the concept "problem" as a component of the scientific method. McDowell (1966) asserts that a problem consists of two elements, an intention and a barrier to the realization of that intention. McDowell says that a problem statement documents the intention, explicates the barrier, *and* presents either data or a convincing logic that establishes the connection between the two.

McDowell's work and the deliberations of the participants in the 1972 NSPER sessions (Gephart, et. al., 1973) make the point that the abstract language of the scientific method is deterrent to the actual work of systematic problem solving. That is, the terms "intention" and "barrier" need to be made more specific *if work* is to be accomplished. To achieve that specificity NSPER: 72 participants delineated three types of intentions[1] and their associated barriers as shown in the table below:

TABLE 1: Intentions and their Barriers

Intention	*Barrier(s)*
1. To know in a generalizable or theory sense.	1a. The item to be known has never been studied before or, 1b. The study or studies of the item are inconclusive.
2. To do some task.	2a. The tools or procedures for doing the task have not been created, or 2b. The tools or procedures for doing the task will not perform at the quantitative or qualitative levels desired.
3. To choose one item (or some subset) from a set for treatment different than the others in the set.	3. The relative worth of each of the items in the choice-making situation is not known.

When we have an intention to know that is problematic, the research process is the appropriate resolution strategy. When done properly the research process produces evidence which supports or rejects conjectures (hypotheses) about bits of phenomena and/or their relationships. The development process, the appropriate resolution strategy when an intention to do is problematic, creates work tested tools or procedures. An intention to choose that is problematic is best resolved by the evaluation process. When done properly, the evaluation process gives information to decision makers – information which helps them determine the relative worth of the alternatives being considered.

---

[1] It is readily recognized that other types of intentions exist. The three dealt with in this paper are commonly misunderstood and mixed together by educators.

Evaluation and development are similar processes in their focus on a specific situation. They are different in that evaluation produces information while development produces tools and procedures. The products of evaluation and research are also similar and different. They both produce information. Research done right produces information that is applicable in different places and times. It is generalizable. Evaluation on the other hand produces situation specific information: What is the relative value of alternatives A, B, C . . . in location Y at time T?

Evaluation then is a systematic problem-solving methodology similar to the research and development processes, but different from them in terms of the specifics of the purpose it serves and what it produces.

THE COMPONENTS OF THE EVALUATION PROCESS

A second way of examining the evaluation process is initiated by the question, "What does it consist of? What are its component parts?" A review of the literature on the evaluation process is indirectly helpful in answering this question. That is, although much of the literature alludes to the components of the evaluation process, few if any writings focus on the delineation of them directly.

Five components can be teased out of the literature (Figure 1). The elimination of any one of them seems logically to emasculate the concept "evaluation."

Those five are:

1) *Need to make a choice*. Every evaluation effort and all of the evaluation models confirm "the need to make a choice" as a central component of the evaluation process. Sometimes this choice making is for simple labeling purposes – we want to

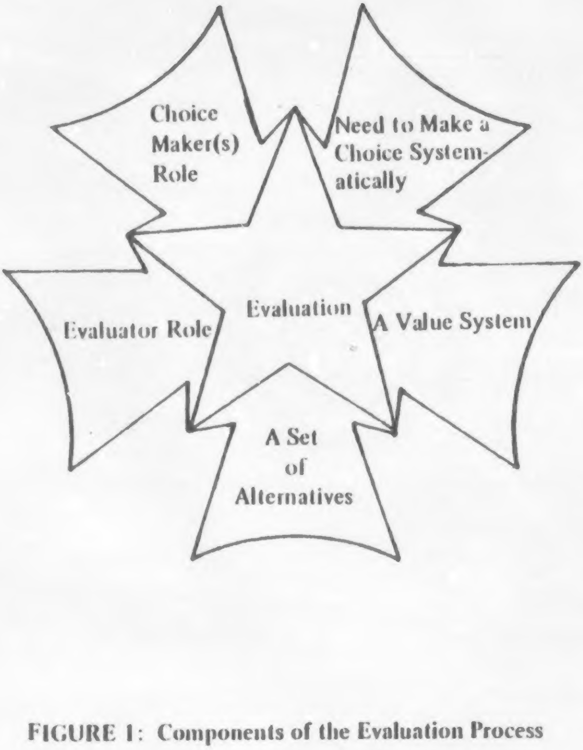

Choice
Maker(s)
Role
Need to Make a
Choice System-
atically
Evaluation
Evaluator Role
A Value System
A Set
of
Alternatives

FIGURE 1: Components of the Evaluation Process

communicate quickly to others in a manner which tells the quality of some entity. The need in this labeling situation stems from the fact that there are a broad range of symbols of quality all of which *cannot* be used simultaneously. For example, in evaluating a student's work a range of symbols from A to E is usually considered, and the evaluator needs to choose the one label that best reflects the quality of the work. At other times the need to make a choice stems from the fact that several action options may exist and it is impossible to simultaneously pursue all of the options. For example, consider the situation in which a new educational program has been through a test run in a school system. At this point the decision makers have several action alternatives: a) build the new program in as a part of the regular program; b) test the program again with some modifications; or c) drop the new program. It is impossible to do all three. Thus there is a need to make a decision.

2) *A value system*. The act of choosing one alternative from some set of alternatives always occurs in a given setting and at a given time. As such, it occurs within a specific value system or value complex. The nature of that value system will establish who makes the decision and the variables to be used in determining the relative worth of the options. In one location a particular decision will be made by the school board members; in another the same decision will be in the hands of administrators; in still others, in the hands of teachers. In one setting the acceptability of the materials by the lay public will be a primary criterion. In another, public opinion will be relatively immaterial. Which (either decision locus or basic criteria) surfaces is an artifact of the extant value base.

3) *A set of alternatives*. Every evaluation problem has within it a set of alternatives. Many fail to differentiate those instances in which information is desired that will *describe* the school's population (on some relevant scale) from those instances in which a choice is to be made. The former is a situation in which we simply want to know, "How do the people of interest distribute on some variable?" (e.g., what are the reading scores of people who successfully leave the sixth grade in school J?) The latter, the choice making, is a situation in which two or more alternatives are present and the circumstances prohibit treating both (or all) of them alike. For example, we could use textbook series A only,

textbook series B only, or some combination. It is impossible to implement *all* three options.

Evaluation always contains or focuses on a set of two or more alternatives. Sometimes only a few alternatives out of some larger set will be considered and used. At other times, a limited subset is readily visible to the decision makers and the process calls for delineation of heretofore unperceived alternatives and then the specification of the relative worth of all the alternatives. If a set of alternatives cannot be delineated, an evaluation problem cannot be delineated.

The distinction between components 1 and 3 needs further comment. At first consideration, listing them separately appears pedantic. There are many instances where the "need to make a choice'" exists within an individual or organization BUT no alternatives exist. At other times options exist but in the absence of a perceived need to choose. Neither of these instances call for an application of evaluation as a formal problem solving strategy.

4) *Evaluator role*. In every instance of an application of the evaluation process there is an evaluator role. Sometimes it is occupied by a person other than the choice maker. At other times one person plays both roles. The evaluator role involves determination of the alternatives that will be considered in the choice making and the value dimensions to be used in determining their relative worth. It also involves generating data and interpreting those data, sometimes in a formal report.

5) *Choice maker role*. The last component is the choice maker role. Again, this may be played by someone other than the player of the evaluator role or one person may play both roles. Further, the choice maker role may be played by one person or by a group of persons. If one person, it may be an instance in which that person is acting independently of others or under the influence of some collection of others. If the decision making role is played by a group, it may be acting as a whole or it may be a group acting serially. In any case, an individual plays the choice maker role by specifying and receiving the information necessary to determine the relative worth of the alternatives and using that information for sorting the alternatives into categories to be treated differently (for example, qualitative descriptors to be used and those not to be used; actions to be taken and those not to be taken, etc.).

A distinction needs to be made at this point between evaluation as a systematic problem solving process and evaluation that is a spur of the moment, accidental activity. The components listed above are found in the instances of evaluation used as a systematic problem solving activity. Evaluation is a continual human activity. Sometimes it is carried out in ways that generate public information. At other times it is used to generate private information. The former is the focus of this presentation.

HOW IT WORKS

Despite the large number of what appears to be competing models of the evaluation process, there are a very limited set of operational definitions of the term. Given acceptance of a degree of abstraction in language it is possible to describe the evaluation process using one set of operations. At a more explicit level of language several process descriptions are possible. The material which follows presents the generalized operational definition and three derivatives that can be shown through a higher degree of specificity in the language used. The fact that differing levels of abstraction (or specificity) of language are being used to describe evaluation operations should be clear in the reader's mind. The different operational descriptions are not different evaluation models, but rather, operational definitions that are more or less abstract.

How It Works – The General Case

The general case of evaluation involves three kinds of activity: Determining the information needed, generating that information, and reporting that information. These three activities can be found in writings about the CIPP Model, the CSE Model, the Discrepancy Model, the Countenance and/or Responsive Models, the Formative and/or Summative Models, the Goal Free Model, etc. The three activities can also be seen in case study reports of evaluation efforts.

Before detailing the general nature of these three classes of evaluation operations, a distinction needs to be made between applications of the evaluation process that are ad hoc in nature and applications that are of a continuing nature. The former are activities undertaken to serve a specific purpose or decision. Their operations are generally sequential.

The operations in a continuing evaluation effort, such as a university's institutional research operations or a public school system's evaluation program are much more iterative in nature. And although this complicates the operations in an actual case, it makes only a simple modification conceptually. That modification is the addition and use of feedback mechanisms between the general activity blocks. (Fig. 2 and 3)

FIGURE 2: A Block Diagram of the Ad Hoc Evaluation Process

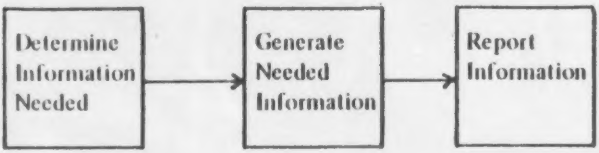
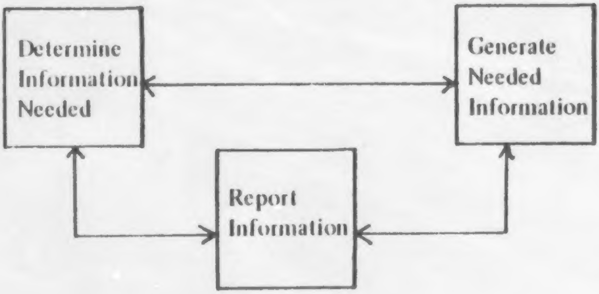
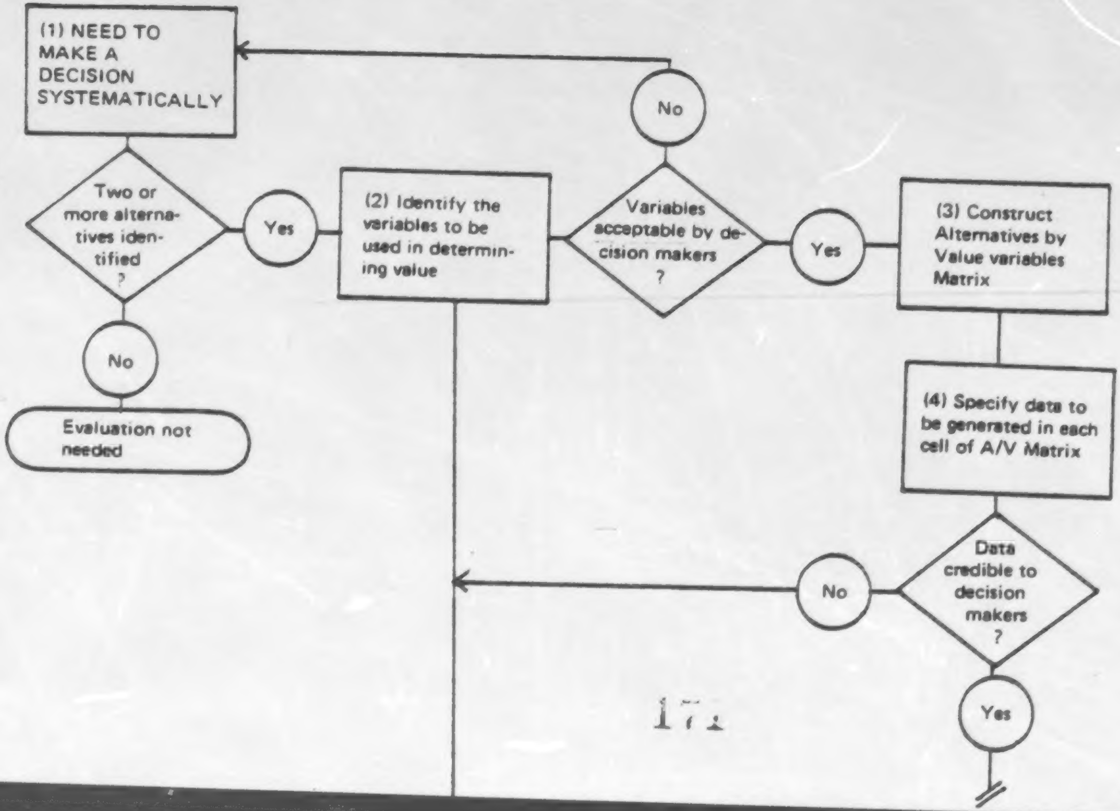

Determine Information Needed
Generate Needed Information
Report Information

FIGURE 3: A Block Diagram of a Continuing Evaluation Operation


Determine Information Needed
Generate Needed Information
Report Information

Determining the information needed is essentially an interface activity in which three items are made explicit. (A flow chart of these activities is presented in Figure 4. First, is there an "evaluation problem?" That is, can a choice making activity be identified in which the relative worth of the alternatives is not known to the individuals who must make the choice? If the answer to that question is no, an evaluation is not needed. Second, a matrix should be constructed which specifies the alternatives and the variables that the decision makers will use in determining the relative worth of the alternatives. An illustrative matrix is shown in Figure 5. In this case the alternatives identified through discussions with the decision maker emanate from a textbook adoption problem in a state that publishes an approved text list. The variables to be used in determining the relative value of those alternatives are identified through discussions with the decision makers and selected educational specialists. The blank column is included to suggest that more variables might be included in the evaluation. The same point should be made about the alternatives. The specifics of the situation and the decision will determine the number and nature of the alternatives and value variables to be included in a given evaluation effort. The third item to be specified is the plan for collecting the data needed to fill each cell of the alternatives/values variables matrix. This involves activities (3), (4), and (5) in the flow chart (Figure 4): Specifying the exact nature of the data; specifying the procedures to be used to generate or collect the data; and, scheduling the data generation activities. In doing these activities, the evaluator needs to check with the decision maker to see that the data are credible; that the procedures to be used are acceptable within policy and budgeting limits; and that the schedule is acceptable.


(1) NEED TO MAKE A DECISION SYSTEMATICALLY
Two or more alternatives identified ?
Yes
No
Evaluation not needed
(2) Identify the variables to be used in determining value
Variables acceptable by decision makers ?
No
Yes
(3) Construct Alternatives by Value variables Matrix
(4) Specify data to be generated in each cell of A/V Matrix
Data credible to decision makers ?
No
Yes

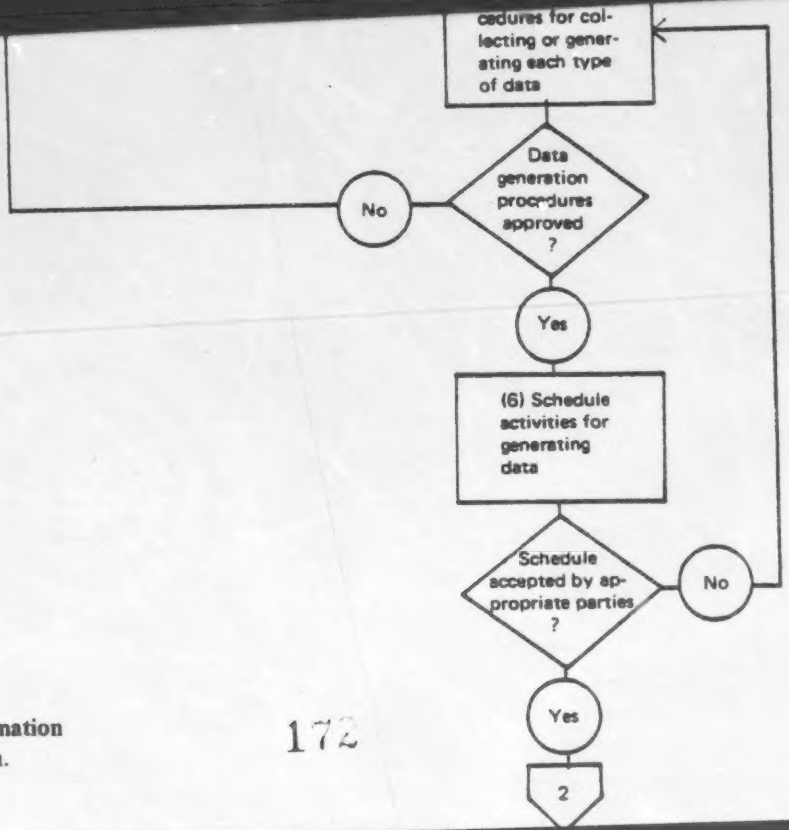

cedures for collecting or generating each type of data
Data generation procedures approved ?
No
Yes
(6) Schedule activities for generating data
Schedule accepted by appropriate parties ?
No
Yes
2

FIGURE 4: A Flowchart of the Information Delineation Aspect of Evaluation.

FIGURE 5: An Illustrative Alternatives by Value Variables Matrix

Alternatives

Textbook Series A	IA	IB	IC	ID	IE	IF
Textbook Series B	IIA	IIB	IIC	IID	IIE	IIF
Textbook Series C	IIIA	IIIB	IIIC	IIID	IIIE	IIIF
Textbook Series D	IVA	IVB	IVC	IVD	IVE	IVF
Textbook Series E	VA	VB	VC	VD	VE	VF
Multiple Texts	VIA	VIB	VIC	VID	VIE	VIF

Cost

Documented Achievement

Acceptability to Staff

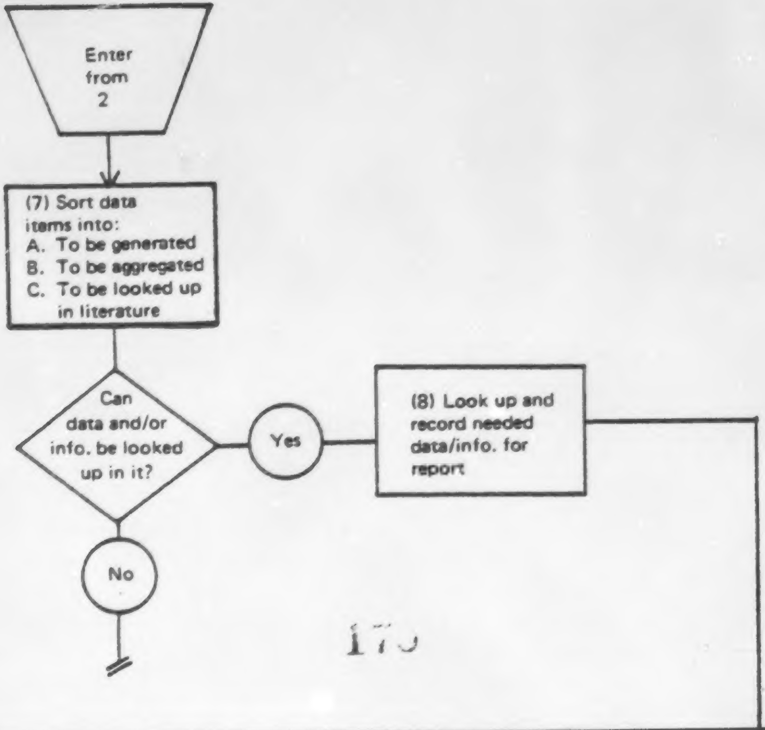
In-Service Education Necessitated

Community Acceptability

Value Variables

The second general activity in the evaluation process is the generation and analysis of data. The purpose of this activity is to produce the data and/or information that is needed to fill the cells of the matrix. In the illustration (Figure 5), cell IA would be filled in by determining the cost for textbook series A (either on a per pupil or total cost basis or both). Cell IIA calls for the comparable cost figure for textbook series B. Cell IB would be filled in with data and/or information on the level of achievement that can be attained through the use of textbook series A. The flowchart presented in Figure 6 displays the activities which comprise this aspect of an evaluation effort. It recognizes that some of the information needed for a decision may already exist in society's knowledge bank; that other kinds of data may be possessed by some other agency (the cost factor in the textbook illustration is a case in point); that still other kinds of decision may exist in the client's data bank but not in a form directly relevant to the decision being served; and finally, that some of the data needed to determine relative value of the alternatives will have to be generated in the evaluation. Associated with each of these classes of information, the flowchart indicates activities that need to be carried out. A destinction is made here between the concepts "data" and "information." The following test scores are data: 25, 31, 16, 42, 36, 34, 29, 34, and 37. Those data must undergo some conversion in order to become information. The ten scores range from 16 to 42, have a mean of 32.0, and a standard deviation of 7.3. Evaluation efforts frequently involve the generation of data that must be converted to information. The flowchart (Figure 6) recognizes two more operational points. All of the information accumulated needs to be interpreted in terms of its meaning for the specific decision. Finally, the flowchart does not delineate between cognitive and affective data. This is deliberate. Usually both types of data will be involved. In some cases affective data will be more central and in others cognitive data items will be preeminant. The exact mix will be determined by the specific project.

The final class of activities in evaluation is reporting. Evaluation theorists take different positions on this activity. Some are content to present the information collected in the evaluative effort. Others believe the evaluative work is incomplete until the


Enter from 2
(7) Sort data items into:
A. To be generated
B. To be aggregated
C. To be looked up in literature
Can data and/or info. be looked up in it?
Yes
(8) Look up and record needed data/info. for report
No

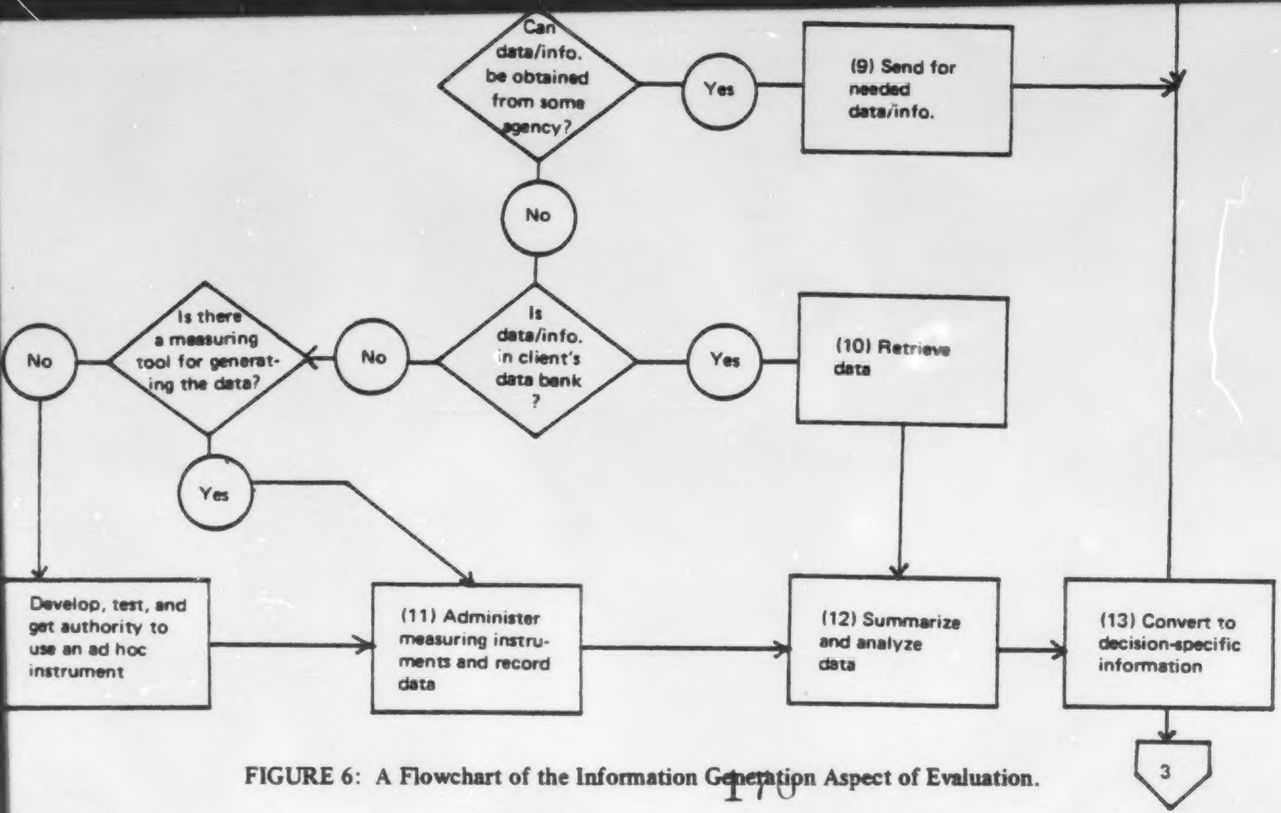
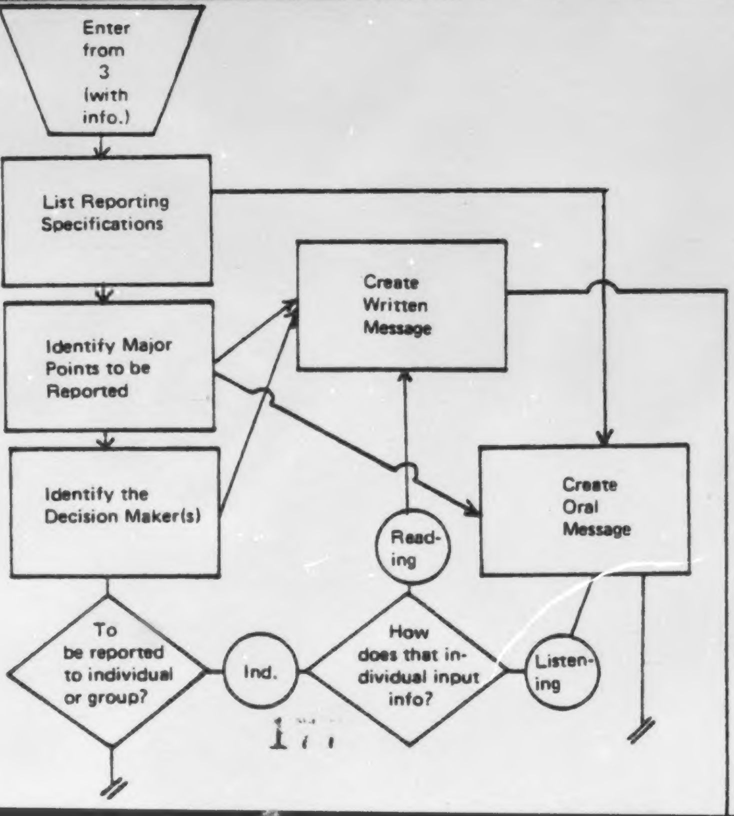

Can data/info. be obtained from some agency?
Yes
(9) Send for needed data/info.
No
Is there a measuring tool for generating the data?
No
Is data/info. in client's data bank ?
Yes
(10) Retrieve data
No
Yes
Develop, test, and get authority to use an ad hoc instrument
(11) Administer measuring instruments and record data
(12) Summarize and analyze data
(13) Convert to decision-specific information
3

FIGURE 6: A Flowchart of the Information Generation Aspect of Evaluation.


Enter from 3 (with info.)
List Reporting Specifications
Identify Major Points to be Reported
Identify the Decision Maker(s)
Create Written Message
Create Oral Message
Read-ing
To be reported to individual or group?
Ind.
How does that individual input info?
Listen-ing

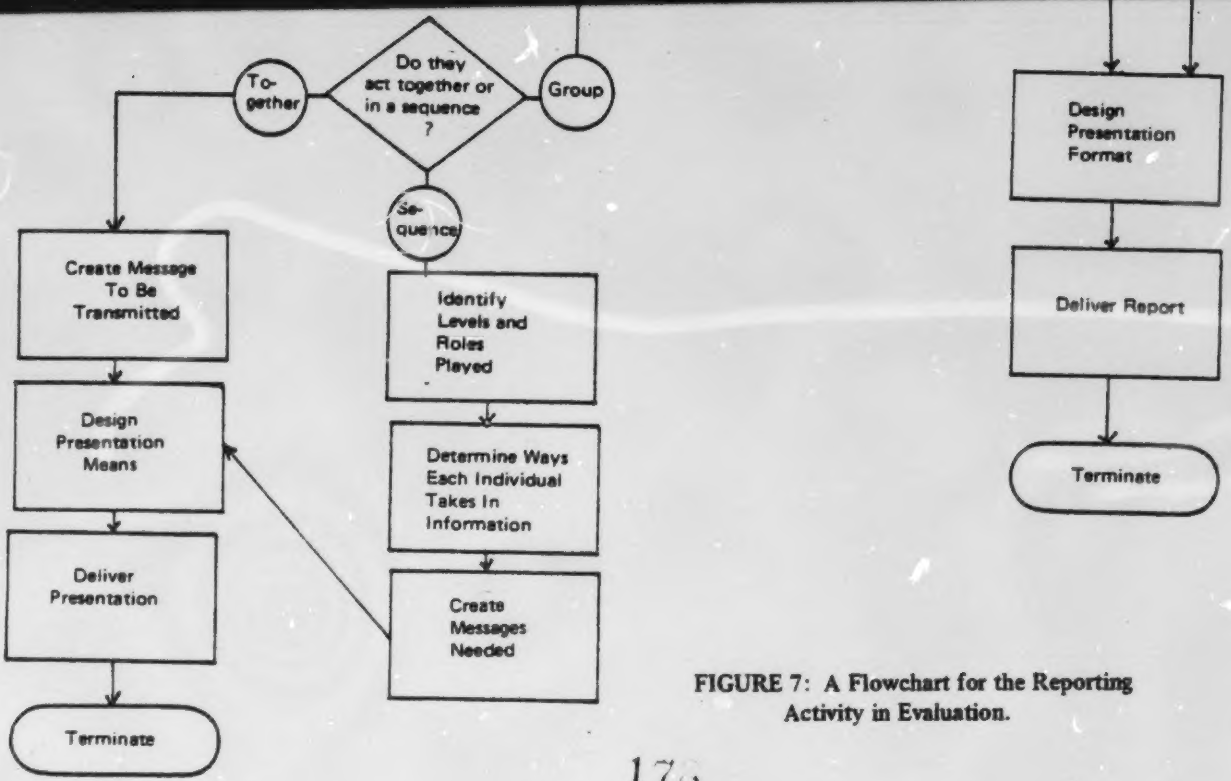

Do they act together or in a sequence ?
To-gether
Group
Se-quence
Create Message To Be Transmitted
Design Presentation Means
Deliver Presentation
Terminate
Identify Levels and Roles Played
Determine Ways Each Individual Takes In Information
Create Messages Needed
Design Presentation Format
Deliver Report
Terminate

FIGURE 7: A Flowchart for the Reporting Activity in Evaluation.

III	Process Components or Activities.	1. Delineation of the typical decisions. 2. Identification of info which might be used in those decisions. 3. Testing & data storage 4. Retrieval and interpretation.	1. Determination of the alternatives that will be considered & criteria to be used in weighing them. 2. Generating the information on all alternatives. 3. Communicating info to decision makers.	1. Identifying competent judge(s). 2. Collecting and analyzing data. 3. Reaching conclusions. 4. Submitting recommendations to authorities.
IV	Products or Outcomes of the Strategy.	Data on characteristics of people or programs encoded in some bank ready to be accessed if useful in a decision situation.	Information which describes the relative worth of a specific set of alternatives in a particular setting.	Statements of worth and/or recommendations for subsequent actions.
V	Criteria Useful in Assessing the Quality of an Evaluation.	1. Fit between data items & typical decisions. 2. Quality of data generation techniques. 3. Retrievability of stored info. 4. Security of info. 5. Cost.	1. Scientific criteria-validity, reliability. 2. Practical criteria-relevance, importance, scope, credibility, timeliness, pervasiveness. 3. Cost.	1. Credentials of judge. 2. Procedures for data generation. 3. Reportage to people in positions able to carry out recommendations.

decision or judgment is made. Regardless of the stance, a reporting effort is included.

Reporting activity, as indicated in Figure 7, starts with the determination of the information needs (Figure 4), for in that aspect some of the reporting specifications are set. As the delineation of the alternatives is carried out some aspects of the reporting activity are set. As the value variables are established, still additional reporting specifications are set. The level of decision making on which the evaluation is to focus further specifies the reporting. The reporting activity involves indentification of the specific individual or individuals who make the choice and the determination of the way in which the person(s) receive or process information. Do they intake information best through reading? Through listening? Through working through illustrative tasks? This information also sets specifications for the reporting activity. Within these specifications evaluation reporting concludes with the design, development, and delivery of the evaluation information.

How It Works: Three General Models

As stated earlier, the operational definition of the evaluation process can be described using language at different levels on a continuum from abstract to specific. The preceding was very abstract, a description of the general case. This section will be more specific by dealing with three general types of evaluation currently described in the evaluation theory literature. These three operational forms of evaluation have been used for decades. The first is the form espoused by those individuals who define evaluation as measurement. The second form is the definition of evaluation as a service to decision making. The third form, by definition, merges measurement and decision making by defining evaluation as a synonym of judgment.

An illustration of the first form (evaluation equals measurement) can be seen in a school system's standardized testing program. The staff is aware of a cyclical need to make certain types of decisions. To help them in making those decisions, a battery of measuring procedures are administered, the responses tabulated, and the data stored ready for retrieval if and when it is needed.

An illustration of the second form can be seen in the work of an evaluation department as it tries to help a school board decide on the set of educational goals their school district should focus on. A large number of goals are possible candidates and the school cannot focus on all of them. The goals to be considered must be delineated and the variables which will be used to determine the relative worth of the many goals must be specified. The data for each goal on each variable are collected, analyzed, and interpreted. And, a report indicating relative worth of the individual goals is made to the group that has the responsibility for the selection of a specific set of goals.

The third form of evaluation (evaluation equals judgment) can be illustrated by the work of an accrediting agency. Knowledgeable judges are selected. Those judges examine the item to be accredited. Their observations are collected and a judgment is rendered. (The evaluation problem in this case is the intention to describe the quality of a program when many descriptors are available and their relative appropriateness is not known.)

These illustrations suggest the operational definitions of the three general forms of evaluation (Figure 8). The figure below item III outlines their operational steps. The other rows detail their general nature, products, and criteria for qualitative assessment of an evaluative application.

In summary, evaluation in any of its forms involves determining the nature of the information needed for a decision; the generation, analysis, and interpretation of data basic to that information; and the reporting of the information. It is a systematic problem-solving strategy useful when there is an intention or need to make a choice and the relative value of the options is not known. Its components are a value base, a set of alternatives, a need to make a choice, and "choice-maker" and "evaluator" roles. It is a natural human behavior engaged in by all. Sometimes it is done subjectively and for the most part privately. At other times it is done systematically and publicly. The latter is our focus here.

THE AFFECTIVE DOMAIN

The presentation of that center heading, "The Affective Domain" in some respects is counterproductive to clear under-

FIGURE 8: The Three General Forms of Evaluation

	Categories of Description	*A* *Perhaps Associated with Decisions* *(Evaluation = Measurement)*	*B* *Isomorphic to a Decision* *(Evaluation = Service to a Decision)*	*C* *Encompassed in a Decision* *(Evaluation = Judgment)*
I	General Nature.	A data banking or measurement process (e.g., standardized test program).	A service to decision making through which relative worth of alternatives is described.	Making judgments re the value of entities and making recommendations.
II	Characteristics of "Problem" for which there is an Appropriate Solution Strategy.	Regularized activities in which decisions may be made which might be facilitated through the use of previously generated data.	1. Intention to choose a specific alternative. 2. The existence of 2 or more possible alternatives. 3. The lack of data that estimates relative worth of all alternatives. 4. Lead time.	1. The existence of some completed educational project or an institution. 2. Competing alternatives. 3. Desire or need to control subsequent operations.

standing. The word, "domain," connotes scope, dominion, power, boundary. It often gives rise to ideas of exclusivity. Educators use the phrase to signal something other than cognitive. In this usage educators have taken an either-or stance, a stance that makes understanding the concept, affect, more difficult because cognition and affectivity are not discrete, not separate in real world behavior.

As a case in point consider the behavior of an artist. Composing and rendering a work of art are prime examples of the inexorable unity of what is called cognitive, affective, and psychomotor behaviors. Logical analyses of behavior have given labels to aspects of behavior FOR THE PURPOSE OF STUDY AND COMMUNICATION about a complex concept. In this manner we have created three useful explanatory fictions, the cognitive domain, the affective domain, the psychomotor domain. BUT, as useful as this separation is, we must never make the mistake that it is real or that we will see "pure" affective behavior or "pure" cognitive behavior, etc.

With this caveat in mind, what is the affective domain? To answer that question, we suggest the ancient "Blind Men and the Elephant" story. To them the definition of "elephant" was varied because of differing vantage points. People describing the affective domain have different vantage points as well. Some describe it by telling its synonyms. Others point to examples of it. Still others will tell you what it consists of. And another group will ply you with how it works. Two more can be found. Some describe by contrasting the affective domain with items that are related but different. And finally, there are people who describe by telling how it fits in the grand scheme.

We agree with Henry Cady (1967): To apprehend the meaning of a complex concept ALL of these forms of description or definition, need to be used. So, to practice that belief, the following discussion merges the message from all of the "blind men."

The affective domain is an abstract categorical term which references a class of concepts. As such, it has no synonyms. In normal and professional discourse, it is used as if it is the opposite (an antonym) of the cognitive domain. This is a mistaken usage for there are cognitive aspects to affective behavior and affective aspects to cognitive behavior.

At times, both in general and professional discourse, a number of terms are used as synonyms (e.g., emotions, attitudes, values, perceptions, anxiety, love, hatred, feelings, beliefs, etc.). There are two difficulties in this. First, the term on which we are focused is "affective domain." The term "emotions" cannot be a synonym of the term "affective domain" for the former, "emotions," as a plural term, references more than one of the entities of interest while the term "affective domain" references a (singular) category. The failure to recognize the movement from a category to an element in that category blocks thorough understanding.

The second difficulty here is shown when people say, "Right! 'Emotions' and the 'affective domain' are not synonyms! 'Emotion' (singular) and 'affect' are synonyms!" Wrong! Our language has a built-in variable, abstract to specific. "Emotion" and "affect" are at different levels of abstraction (or specificity) just as "nephew" and "aunt" are at different levels on a kinship variable. "Affect" is a more inclusive concept than is "emotion." And, "the affective domain" is a still higher level of abstraction.

During the NSPER: 76 sessions, note was made of the fact that the paper presenters and the participants used over a dozen words in talking about the focus of interest, the affective domain. In some of the discussion there was at least the hint that these terms (and concepts) comprise a taxonomy. That is, one of the terms, "perception" encompassed another, "self-perception." In turn "perception" was encompassed by some other term. The groups at the three NSPER: 76 sessions were asked by this writer to help construct a taxonomy of this terminology. An effort of this sort was made. And, although the taxonomy that resulted was far from definitive, it did seem to help people apprehend meaning of the concept, "the affective domain." For that reason, and to spur reaction, that tentative taxonomy is presented in Figure 9.

The affective domain is the term for the whole class and as such is the most abstract concept in the taxonomy. Comments by NSPER: 76 speakers suggest that the affective domain has two subdivisions, the physiological and the psycho-social. It was further suggested that the former was more often the concern of

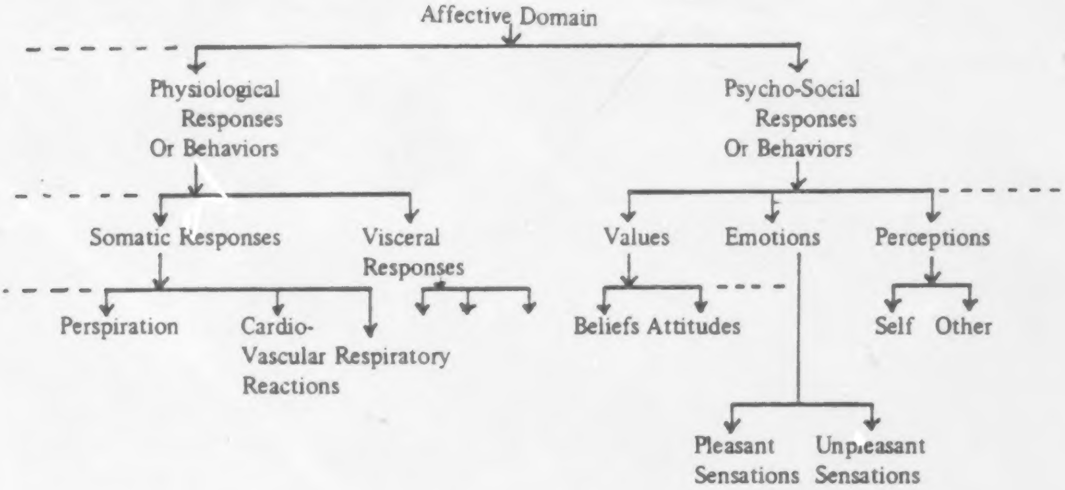

Affective Domain
Physiological Responses Or Behaviors
Psycho-Social Responses Or Behaviors
Somatic Responses
Visceral Responses
Values
Emotions
Perceptions
Perspiration
Cardio-Vascular
Respiratory Reactions
Beliefs
Attitudes
Self
Other
Pleasant Sensations
Unpleasant Sensations

FIGURE 9: Toward a Taxonomy of the Affective Domain

the medical field and the latter of the fields of education, psychology, and sociology. The dotted lines in Figure 9 suggest that additional categories might exist.

The physiological class, according to knowledge of our nervous system, can be subdivided into somatic and visceral responses. The former includes those internal changes in body temperature, respiration, pulse rate, and blood pressure. The latter are the sensors for external stimuli and muscle contractions and expansion.

The other major branch of the affective domain subdivides into categories labeled "values," "emotions," and "perceptions." Values, in the NSPER: 76 discussions, seemed to encompass beliefs and attitudes; emotions divide into pleasant and unpleasant sensations; and perceptions into those of self and others.

The reader should keep in mind that the classification in Figure 9 is proposed only as a suggested direction. The NSPER: 76 participants and this writer have serious reservations about the validity of the placement of the terms. Despite the caveat, three things can be accepted from this effort to "taxonomize." First, the domain does have terminology at differing levels of abstraction (or specificity). Second, a taxonomy of affective domain terms seems to have potential for enhancing our understanding. Third, in the affective domain involves measurement, and thus, focuses on the specific. Measurement of gross abstractions are very difficult and usually accomplished by combining observations of more specific things. Therefore, measurement in the affective domain is likely to focus on perceptions of self and other, on attitudes and beliefs, and on pleasantness. Again, although they encourage further analysis along these lines, the participants in the NSPER: 76 sessions do not endorse the particulars of the classification system presented in Figure 9.

To this point our definition effort has dealt with synonyms and antonyms. In so doing however, we have touched on constitutive definition – what it consists of. The discussion indicates that the components of the affective domain are concepts of varying concreteness or substance. Further, there is confusion in the field about them. That confusion is demonstrated when persons try to state the inclusion-exclusion criteria that would determine whether concept X is or is not included in the taxonomy or domain.

Confusion is also apparent when we try to construct an operational definition – how does it work? The effort by Krathwohl and others on a taxonomy of the affective domain, although our best effort to date, is more an operational definition than it is a taxonomy. Its levels start with receiving and move through responding, valuing, organizing, and characterizing. Such terminology classifies levels of operations.

We can define by giving illustrations or pointing to instances in which affect exists. Ostensive definition is possibly our strongest suit when it comes to defining affect. Unfortunately, such a definitional approach is not accorded much prestige. Our culture does not recognize ostensive definition as a very important mode of definition. We want to count, measure, and operationalize, a point that will be picked up in the final section of this paper.

Defining the affective domain by contrasting it with related items is another approach that needs attention. The "affective domain" should be contrasted with "affect," with the "cognitive domain" the "psychomotor domain," and other related concepts. This writer is incapable of making these contrasts. Someone with greater knowledge in this subject matter is needed to do the job.

The importance of this form of definition can be seen by analogy. Consider for a moment a contrast definition of reading.

> Reading is like listening, writing, and speaking. They all involve messages transmitted from person to person. Listening and reading focus on the receiving of coded messages. They are different in that reading deals with messages encoded in a visual media while the messages in listening are encoded in an aural media. Reading is like writing. They both involve visually coded message processing. They differ in that reading is decoding and writing is encoding.

Work of this sort on the "affective domain" is needed to help us apprehend the meaning of the concept.

The last form of definition can be called classificatory definition – where does it fit in the grand scheme of things? We have rudimentary work here. The affective domain is an aspect of behavior akin to the cognitive and psychomotor domains. As such they assist us in analyzing behavior psychologically and sociologically, that is, what is going on within and between individuals.

As stated above, our attempts to apprehend the meaning of the affective domain are complicated due to fragmentation of previous work and our failure to attend to some forms of definition. Our understanding would be enhanced if: (1) we developed and carried out systematic definition efforts on those forms of definition that have been neglected to date; and (2) we find a way of synthesizing the meaning presented from all the definition forms. The whole is more than the sum of its parts.

ACCEPTANCE OF OUR AFFECTIVE EPISTEMOLOGY

One of the barriers to knowledge and utilization of the affective domain in evaluation can be found in the epistemology we believe to be preeminant. How do we know that we know something? In our culture (the Euro-American heritage) we know that we know something when we can classify it and count and measure it. *And*, we really know that thing when all of us get the same numbers! This epistemology is reinforced over and over by schools, industry, government, social press for accountability, etc: "Let's test it," "Measure it so we can know what's happening."

There is another epistemology, a primarily affective one, that has been attributed to the Afro-American heritage. In that epistemology we know that we know something when we can feel it. As implied earlier, this knowing-by-feeling epistemology is not accorded the prestige that is given to the count-and-measure approach. But, it is a way of knowing that we ought to recognize *and* use. Further, the count-and-measure epistemology is not free from error. In fact there are classes of error to which we become vulnerable when we rely on it for knowing, errors of logic, of measurement, and of analysis.

In our recent history there are many examples of our rejection of the feeling epistemology. Title I of the Elementary and Secondary Education Act provided an abundance of them. That act mandated that the projects funded through it be evaluated. To the education community, and others, that meant, "count it and measure it!" So we did. Time after time the *numbers* we got failed to show any differences in effect. So we dropped those procedures. Many times that termination caused bad feelings. The people in those programs were excited about what was happening. They could *feel* a change! But that wasn't good enough! Knowing by feeling was no match for knowing by the numbers!

Many people strongly believe that feeling is an unreliable way of knowing. But, how do they know that? Mostly by feeling! To the writer it is sheer folly to reject any way of knowing! The person addicted to a single epistemology, who denies himself the use of all other epistemologies, is much more likely to err than is the person who uses various epistemologies as checks and balances on each other.

The message is clear. If we are to evaluate affective aspects of educational programing, we must accept and become more skilled at sensing, at feeling what is going on. That statement *does not* mean that we should stop counting and measuring. Rather, it is born out of a recognition that our attitudes about how to really know something too often exclude the use of affectivity. And, it is born out of a personal recognition of how tough it is to get in touch with my own feelings! It is something I must continually work on.

SUMMARY

The affective domain relates to evaluation in two ways: as an aspect of programs or products to be evaluation, and as a vehicle for the evaluation. To use this relation requires knowledge of the evaluation process and of the meaning of the concept, the affective domain.

Evaluation is described here as a systematic problem-solving process appropriate for situations in which choice making is our intention and we do not know the relative worth of the options that structure the choice situation. Evaluation is described operationally and componentially.

The affective domain references an aspect of behavior that has been isolated via logical analysis from the remaining aspects of behavior. Our knowledge base related to the concept, affect, is sparce. To help suggest directions of investigation six definitional approaches are listed and applied and a tentative taxonomy of terminology is presented.

Finally, it is argued that affect as an epistemology deserves greater acceptance; that the use of a single epistemology (usually the cognitive) is more prone to error than the use of several; and, that work is needed to help us get in touch with our feelings and thus improve our ability to know via affectivity. Try it! It feels good!!!

REFERENCES

Cady, H. Toward a definition of music education. In H. Cady (ed.) *A conference on research in music education*. Cooperative Research Project No. 6-1388. Columbus, Ohio: The Ohio State University Research Foundation, 1967. (ERIC Document ED 013 973)

Gephart, W. J., Ingle, R. B. and Saretsky, G. *Similarities and differences in the research and evaluation process.* Manuscript developed from discussions and follow-up of NSPER: 72, Bloomington, Ind.: Phi Delta Kappa, 1973.

McDowell, E. C. *A definition of "problem."* Unpublished masters thesis, University of Chicago, 1966.